



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 98347

To: Daniel Sullivan  
Location: CM1-11E12  
Art Unit: 1636  
Thursday, July 17, 2003

Case Serial Number: 09/914191

From: Beverly Shears  
Location: Biotech-Chem Library  
CM1-1E05  
Phone: 308-4994

beverly.shears@uspto.gov

### Search Notes

Daniel,

Due to seq. length, unable to search amino acid query. I did, however, search Seq. ID 1 in the AA dBs.

Beverly

Note: Qy has an N in seq so 99.8% = 100% match

**THIS PAGE BLANK (USPTO)**

From: Sullivan, Daniel  
Sent: Tuesday, July 08, 2003 4:18 PM  
To: STIC-Biotech/ChemLib  
Subject: Search Request 09914191

Please search for the following in the commercial and issued patent databases:

-a nucleic acid comprising SEQ ID NO:1

Also, I have reason to believe that the reverse complement of SEQ ID NO:1 encodes all or a portion of the following amino acid sequence:

MTHVASQFASSYVIFYWRDYFEDQPLL YPPGFDGRVVVYPSNQTLKDYLSWRQADCHINNLYNTVFW  
ALIQQSGLTPVQAQGR LQGTLAADKNEILFSEFNINYNNELPMYRKGT VLIWQKVDEVMTKEIKLPTE  
MEGKKMAVTRTRTKPVPLHCDIIGDAFWKEHPEILDEDS

would it be possible to search this amino acid sequence against the commercial and issued patent databases as well?

Thanks,  
Dan

Daniel M. Sullivan  
Examiner AU 1636  
Room: 12D12  
Mail Box: 11E12  
Tel: 703-305-4448

RECEIVED  
JUL - 8 2003  
(STIC)

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**



# SEARCH REQUEST FORM

Requestor's  
Name: \_\_\_\_\_

Serial  
Number: \_\_\_\_\_

Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

## STAFF USE ONLY

Date completed: 07-16-03

Searcher: Beverly 24994

Terminal time: 20

Elapsed time: \_\_\_\_\_

CPU time: \_\_\_\_\_

Total time: 25

Number of Searches: \_\_\_\_\_

Number of Databases: 1

### Search Site

☐ STIC

☐ CM-1

☐ Pre-S

### Type of Search

☐ N.A. Sequence

☐ A.A. Sequence

☐ Structure

☐ Bibliographic

### Vendors

☐ IG

☐ STN

☐ Dialog

☐ APS

☐ Geninfo

☐ SDC

☐ DARC/Questel

☒ Other CGN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 9, 2003, 12:44:53 ; Search time 38.5 Seconds  
(without alignments)  
4139.422 Million cell updates/sec

Title: US-09-914-191-1  
Perfect score: 1096  
Sequence: 1 ttggaatagttcttcttta.....gggttagtcagattgttg 598

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-O=/cgr21/USPTO.spool/US09914191/runat.09072003.112348.12858/app.query.fasta.1.775  
-DB=A\_Geneseq\_101002 -QMT=FASTAN -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LIST=45 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-MODE=LOCAL -OUTFMT=ptc -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15  
-USBR=US09914191 @CGN 1.114 @runat.09072003.112348.12858 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_101002.\*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	544	49.2	173	22	AAB94556	Human protein sequ
C 2	544	49.2	298	22	AAM38811	Human polypeptide
C 3	537	48.6	298	22	AAB93770	Human protein sequ
C 4	455	41.2	85	22	AAG74471	Human colon cancer
C 5	427	38.6	297	22	AAM40597	Human polypeptide
C 6	219	19.8	286	22	ABB60415	Drosophila melanog
C 7	129.5	11.7	497	21	ABG39161	Arabidopsis thalia
C 8	129.5	11.7	502	21	ABG39160	Arabidopsis thalia
C 9	129.5	11.7	520	21	ABG39159	Arabidopsis thalia
C 10	128	11.6	470	21	AAG48957	Arabidopsis thalia
C 11	128	11.6	505	21	AAG48956	Arabidopsis thalia
C 12	128	11.6	523	21	AAG48955	Arabidopsis thalia
C 13	90.5	8.3	694	23	ABE92421	Herbicidally activ
C 14	88.5	8.1	128	21	AAG58628	Arabidopsis thalia
C 15	81.5	7.4	982	19	AAW7288	Zebraphish differ
C 16	80	7.3	233	21	AB56951	Human prostate can
C 17	80	7.3	673	21	AA96449	Human prostate can
C 18	79	7.2	224	22	ABG03256	Novel human diagno
C 19	78	7.1	743	22	ABG09442	Novel human diagno
C 20	78	7.1	833	23	AAW47757	Mature EAI protein
C 21	77.5	7.0	277	20	AA937752	Amino acid sequenc
C 22	76	6.9	419	22	AG98793	Human cell death p
C 23	76	6.9	482	22	AAU27892	Human contig poly
C 24	76	6.9	488	22	AG98787	Human cell death p
C 25	76	6.9	490	21	AA949289	Mouse GLC1A polype
C 26	76	6.9	490	22	AG98786	Human cell death p
C 27	76	6.9	578	22	AAU27720	Human full-length
C 28	76	6.9	580	21	AA93274	Amino acid sequenc
C 29	76	6.9	625	22	AB95823	Human protein sequ
C 30	76	6.9	1167	22	ABE04310	Human cardiac aden
C 31	75.5	6.8	393	17	AAW04404	Human CRTAM. Homo
C 32	75.5	6.9	604	23	ABP41697	Human ovarian anti
C 33	75.5	6.9	987	22	AA939227	Human polypeptide
C 34	75.5	6.9	1086	22	AAW41013	Human polypeptide
C 35	75	6.8	227	22	ABG14493	Novel human diagno
C 36	75	6.8	393	21	AA957000	Human prostate can
C 37	75	6.8	1195	22	AAU43884	Propionibacterium
C 38	75	6.8	1583	22	ABG09233	Novel human diagno
C 39	75	6.8	1583	22	ABG30211	Novel human diagno
C 40	74.5	6.7	333	23	ABG61771	Novel cathepsin-L
C 41	74.5	6.8	416	22	AAU32725	Novel human secret
C 42	74.5	6.8	505	22	ABE71148	Drosophila melanog
C 43	74.5	6.7	750	21	AA92831	Mutant human prost
C 44	74	6.7	202	12	AA910744	Non-A non-B hepati
C 45	74	6.8	240	21	AA927963	Arabidopsis thalia

ALIGNMENTS

RESULT 1  
AAB94556  
ID AAB94556 standard; Protein; 173 AA.  
XX  
AC AAB94556;  
XX  
DT 26-JUN-2001 (first entry)  
XX  
DE Human protein sequence SEQ ID NO:15324.  
XX  
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX  
OS Homo sapiens.  
XX  
PN EP1074617-A2.  
XX  
PD 07-FEB-2001.  
XX  
PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX PA (HELI-) HELIX RES INST.  
 XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX DR WPI; 2001-318749/34.  
 XX PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX PS Claim 8; SEQ ID 15324; 2537pp + CD ROM; English.  
 XX CC The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX SQ Sequence 173 AA;

Alignment Scores:  
 Pred. No.: 4, 86e-52 Length: 173  
 Score: 544.00 Matches: 103  
 Percent Similarity: 99.04% Conservative: 0  
 Best Local Similarity: 99.04% Mismatches: 1  
 Query Match: 49.23% Indels: 0  
 DB: 22 Gaps: 0

US-09-914-191-1 (1-598) x AAB94556 (1-173)

QY 598 CACATCTGGACTTAACACAGTACAGCCAGGAGGATTACAGGAACTCTTCAGCA 539  
 DB 70 GlnGlnSerGlyLeuThrProValGlnAlaGlnGlyArgLeuGlnGlyThrLeuAlaA 89  
 QY 538 GACAAGAAATCAGATTTGTTTCTCGAATTCACATCAACTATATATAGTGCCTGATG 479  
 DB 90 AspLysAsnGluLeuLeuPheSerGluPheAsnIleAsnTyrAsnAsnGluLeuProMet 109  
 QY 478 TATAGGAAGGGACTGTGTTGATATATGCGAAGAGTGGATGATGACAAAGAAAT 419  
 DB 110 TyrArgLysGlyThrValLeuIleTgPInLysValAspGluValMetThrLysGluIle 129  
 QY 418 AAGCTGCCACAGAAATGGAAGAAACAGATGCGACTGACCCGACCGACGACCAAGCC 359  
 DB 130 LysLeuProThrGluMetGluGlyLysLysMetAlaValThrArgThrArgThrLysPro 149  
 QY 358 GTGCCCTTGACATGCGATATATCATCGGGGATGCTTTCTTGAAGGAACATCCAGAGATTCTA 299

Db 150 ValProLeuHisCysAspIlelleGlyAspAlaPheTrpLysGluHisProGluIleLeu 169  
 QY 298 GATGAAGACAGC 287  
 Db 170 AspGluAspSer 173  
 RESULT 2  
 AAM38811  
 ID AAM38811 standard; Protein; 298 AA.  
 XX AC AAM38811;  
 XX DT 22-OCT-2001 (first entry)  
 XX DE Human polypeptide SEQ ID NO 1956.  
 XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia.  
 XX OS Homo sapiens.  
 XX PN WO200153312-A1.  
 XX PD 26-JUL-2001.  
 XX PF 26-DEC-2000; 2000WO-US34263.  
 XX PR 21-JAN-2000; 2000US-0488725.  
 XX PR 25-APR-2000; 2000US-0552317.  
 XX PR 09-JUL-2000; 2000US-0598042.  
 XX PR 19-JUL-2000; 2000US-0620312.  
 XX PR 03-AUG-2000; 2000US-0653450.  
 XX PR 14-SEP-2000; 2000US-0662191.  
 XX PR 19-OCT-2000; 2000US-0693036.  
 XX PR 29-NOV-2000; 2000US-0727344.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX WPI; 2001-442253/47.  
 XX N-PSDB; AAI57967.  
 XX PT Novel nucleic acids and polypeptides, useful for treating disorders  
 XX such as central nervous system injuries -  
 XX PS Example 3; SEQ ID NO 1956; 10078pp; English.  
 XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
 CC the encoded polypeptides (AAM38642-AAM4213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activ/inhibit activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: the sequence data for this patent did not form part of the printed  
 CC specification.  
 XX SQ Sequence 298 AA;

## Alignment Scores:

Pred. No.: 5,7e-52 Length: 298  
 Score: 544.00 Matches: 103  
 Percent Similarity: 99.04% Conservative: 0  
 Best Local Similarity: 99.04% Mismatches: 1  
 Query Match: 49.23% Indels: 0  
 DB: 22 Gaps: 0

US-09-914-191-1 (1-598) x AAB38811 (1-298)

QY 598 CACAACTCTGACTACACAGTACAGCCCAAGGGAGATTACAGGAACTCTTNCAGCA 539  
 |||||  
 Db 195 GInGInSerGlyLeuThrProValGInAlaGInGlyArgLeuGInGlyThrLeuAlaA 214  
 |||||  
 QY 538 GACAAGAATCAGATTGTTCTTGAATTCACATCAACTATAATATGATGAGTCCGATG 479  
 |||||  
 Db 215 AspLysAsnGluLeuPheSerGluPheAsnIleAsnTyrAsnAsnGluLeuProMet 234  
 |||||  
 QY 478 TATAGGAAGGAGCTGTGTTGATATGCGAGAAGGTGGATGAAGTATGACAAAGAAATT 419  
 |||||  
 Db 235 TyrArgLysGlyThrValLeuLeuTyrGlnLysValAspGluValMetThrLysGluLe 254  
 |||||  
 QY 418 AACTGCGCAACAGAAATGGAAGGAAAAAAGATGCAGTGCACCGGACCGAGCAAGCCCA 359  
 |||||  
 Db 255 LysLeuProThrGluMetGluGlyLysLysMetAlaValThrArgThrLysPro 274  
 |||||  
 QY 358 GTCCCTTGCACGTCGATATCATCGGGGATGCTTCTTGAAGGAACATCCAGAGATTCTA 299  
 |||||  
 Db 275 ValProLeuHisCysAspIleIleGlyAspAlaPheTrpLysGluHisProGluLeuLe 294  
 |||||  
 QY 298 GATGAAGACAGC 287  
 |||||  
 Db 295 AspGluAspSer 298

## RESULT 3

AAB93770

ID AAB93770 standard; Protein; 298 AA.

AC AAB93770;

DT 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:13475.

DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

OS EP1074617-A2.

PN 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -

XX

PS Claim 8; SEQ ID 13475; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 298 AA;

## Alignment Scores:

Pred. No.: 3,46e-51 Length: 298  
 Score: 537.00 Matches: 102  
 Percent Similarity: 98.08% Conservative: 0  
 Best Local Similarity: 98.08% Mismatches: 2  
 Query Match: 48.60% Indels: 0  
 DB: 22 Gaps: 0

US-09-914-191-1 (1-598) x AAB93770 (1-298)

QY 598 CACAACTCTGACTACACAGTACAGCCCAAGGGAGATTACAGGAACTCTTNCAGCA 539

Db 195 GInGInSerGlyLeuThrProValGInAlaGInGlyArgLeuGInGlyThrLeuAlaA 214

QY 538 GACAAGAATCAGATTGTTCTTGAATTCACATCAACTATAATATGATGAGTCCGATG 479

Db 215 AspLysAsnGluLeuPheSerGluPheAsnIleAsnTyrAsnAsnGluProMet 234

QY 478 TATAGGAAGGAGCTGTGTTGATATGCGAGAAGGTGGATGAAGTATGACAAAGAAATT 419

Db 235 TyrArgLysGlyThrValLeuLeuTyrGlnLysValAspGluValMetThrLysGluLe 254

QY 418 AACTGCGCAACAGAAATGGAAGGAAAAAAGATGCGAGTGCACCGGACCGAGCAAGCCCA 359

Db 255 LysLeuProThrGluMetGluGlyLysLysMetAlaValThrArgThrArgThrLysPro 274

QY 358 GTCCCTTGCACGTCGATATCATCGGGGATGCTTCTTGAAGGAACATCCAGAGATTCTA 299

Db 275 ValProLeuHisCysAspIleIleGlyAspAlaPheTrpLysGluHisProGluLeuLe 294

QY 298 GATGAAGACAGC 287

Db 295 AspGluAspSer 298

## RESULT 4

AAG74471

ID AAG74471 standard; Protein; 85 AA.

XX AAG74471;

AC AAG74471;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:5235.

DE Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX colorectal carcinoma.

KW

XX OS Homo sapiens.  
XX PN WO200122920-A2.  
XX PD 05-APR-2001.  
XX PF 28-SEP-2000; 2000WO-US26524.  
XX PR 29-SEP-1999; 99US-0157137.  
XX PR 03-NOV-1999; 99US-0163280.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
XX WPI; 2001-235357/24.  
XX DR N-PSDB; AAH33902.  
XX PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
XX useful for preventing, diagnosing and/or treating colorectal cancers -  
XX Claim 11; Page 6911-6912; 9803pp; English.  
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
XX cancer-associated nucleic acid molecules (N) and proteins (P), where  
XX the proteins are collectively known as colon cancer antigens. The colon  
XX cancer antigens have cytostatic activity and can be used in gene  
XX therapy and vaccine production. N and P may be used in the prevention,  
XX diagnosis and treatment of diseases associated with inappropriate P  
XX expression. For example, N and P may be used to treat disorders  
XX associated with decreased expression by rectifying mutations or deletions  
XX in a patient's genome that affect the activity of P by expressing  
XX inactive proteins or to supplement the patients own production of P.  
XX Additionally, N may be used to produce the colon cancer-associated P,  
XX by inserting the nucleic acids into a host cell and culturing the cell  
XX to express the proteins. N and P can be used in the prevention, diagnosis  
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
XX and AAB7789 represent sequences used in the exemplification of the  
XX present invention.  
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
XX missing at time of publication, meaning no sequences are present for  
XX SEQ ID NO:1027 to 1052, 7921 and 7922.  
XX SQ Sequence 85 AA;  
  
Alignment Scores:  
Pred. No.: 3,55e-42 Length: 85  
Score: 455.00 Matches: 85  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 41.18% Indels: 0  
DB: 22 Gaps: 0  
  
US-09-914-191-1 (1-598) x AAG74471 (1-85)  
QY 541 GCAGACAGATGAGATTTGTTCTGATTCACATCACTAATATGATGCTCCG 482  
DB 1 AlaAspLysAsnGluLeuLeuPheSerGluPheAsnIleAsnTyrAsnAsnGluLeuPro 20  
QY 481 ATGTATAGGAAGGGACTGTGTTGATATGCGACAGGTGATGATGATGACAAAGAA 422  
DB 21 MetTyrArgLysGlyThrValLeuIleTrpGlnLysValAspGluValMetThrLysGlu 40  
QY 421 ATTAAGCTGCCACAGAAATGGAAGGAAAAAGATGCGAGTGCACCGGACCGAGCAAG 362  
DB 41 IleLysLeuProThrGluMetGluGlyLysLysMetAlaValThrArgThrArgThrLys 60  
QY 361 CCAGTCCCTTCACATCGCATATCATCGGGATGCTTCTGGAGGACATCATCAGAGATT 302  
DB 61 ProValProLeuHisCysAspIleIleGlyAspAlaPheTrpLysGluHisProGluIle 80  
QY 301 CTAGATGACAGACAGC 287

Db 81 LeuAspGluAspSer 85  
RESULT 5  
AAM40597  
ID AAM40597 standard; Protein; 297 AA.  
XX AC AAM40597;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polypeptide SEQ ID NO 5528.  
XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX OS Homo sapiens.  
XX PN WO200153312-A1.  
XX PD 26-JUL-2001.  
XX PF 26-DEC-2000; 2000WO-US34263.  
XX PR 21-JAN-2000; 2000US-0488725.  
XX PR 25-APR-2000; 2000US-0552317.  
XX PR 09-JUL-2000; 2000US-0598042.  
XX PR 19-JUL-2000; 2000US-0620312.  
XX PR 03-AUG-2000; 2000US-0653450.  
XX PR 14-SEP-2000; 2000US-0662191.  
XX PR 19-OCT-2000; 2000US-0693036.  
XX PR 29-NOV-2000; 2000US-0727344.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;  
PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
XX DR N-PSDB; AAI59753.  
XX PT Novel nucleic acids and polypeptides, useful for treating disorders  
XX such as central nervous system injuries -  
XX Example 2; SEQ ID NO 5528; 10078pp; English.  
XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and  
XX the encoded polypeptides (AAM38642-AA42213) with nootropic,  
XX immunosuppressant and cytostatic activity. The polynucleotides are useful  
XX in gene therapy. A composition containing a polypeptide or polynucleotide  
XX of the invention may be used to treat diseases of the peripheral nervous  
XX system, such as peripheral nervous injuries, peripheral neuropathy and  
XX localised neuropathies and central nervous system diseases, such as  
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
XX utilisation of the activities such as: Immune system suppression,  
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening  
XX assays for receptor activity, arthritis and inflammation, leukaemias and  
XX C.N.S disorders.  
XX Note: The sequence data for this patent did not form part of the printed  
XX specification.  
SQ Sequence 297 AA;  
  
Alignment Scores:  
Pred. No.: 6,94e-39 Length: 297  
Score: 427.00 Matches: 89

Percent Similarity: 67.42% Conservative: 0  
Best Local Similarity: 67.42% Mismatches: 5  
Query Match: 38.64% Indels: 38  
DB: 22 Gaps: 2

US-09-914-191-1 (1-598) x AAM40597 (1-297)

QY 598 CACAATCTGGACTAACACAGTACAGCCAGGGAGATTACAGGGAATCTTNCAGCA 539  
DB 203 GlnGlnSerGlyLeuThrProValGlnAlaGlnGlyArgLeuGlnGlyThrLeuAlaA 222  
QY 538 GACAGAGATGAGATTTGTTTCTGCAATTCACATCACTATAATATGAGCTGCCGATG 479  
DB 223 AspLysAsnGluLeuLeuPheSerGluPheAsnIleAsnTyrAsnAsnGluProMet 242  
QY 478 TATAGGAAGAGGACTGTGTTGATATGCAGAGAGTGATGAAGTGAATGACAAAGAAAT 419  
DB 243 TyrArgLysGlyThrValLeuLeuTyrGlnLysValAspGluValMetThrLysGlu 262  
QY 418 AAGCTGCCAACAGAAATGGAAGGAAAAAGATGCGACTGACCCGGACAGGACAAAGCCA 359  
DB 263 LysLeuProThrGluMetGluGlyLysLysMetAlaValThrArgThrArgThrLysPro 282  
QY 358 GTG---CCCTTGCACTCGATATCATCGGGATGCTTCTGGAAGGAACATCCAGAT 302  
DB 283 CysLysProSerHisLeu----- 288  
QY 301 CTAGATGAAGACAGCTGACCCCTTTGGCGCTTCAGTTCTGGTGTCTTAACCATGCAAGCC 242  
DB 288 ----- 288  
QY 241 CTCCCACTCCAGGGCTCCTTGCTTGTGCTGAGTGGTGTG 206  
DB 289 -----ProArgAlaProCysLeuArgTrpLeu 297

## RESULT 6

ABB60415  
ID ABB60415 standard; Protein; 286 AA.  
AC ABB60415;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 8037.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WC200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL04518.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
XX Disclosure; SEQ ID NO 8037; 21pp + Sequence Listing; English.  
PS  
XX The invention relates to an isolated nucleic acid detection reagent  
CC

CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 286 AA;

Alignment Scores: 1.26e-15 Length: 286  
Pred. No.: 219.00 Matches: 47  
Percent Similarity: 60.61% Conservative: 13  
Best Local Similarity: 47.47% Mismatches: 21  
Query Match: 19.82% Indels: 18  
DB: 22 Gaps: 2

US-09-914-191-1 (1-598) x ABB60415 (1-286)

QY 595 CAATCTGGACTAACACAGTACAGCCAGGGAGATTACAGGGAATCTTNCAGCAGAC 536  
DB 166 GluLysGlyLeuThrAsnGlnGlnAlaGluAlaLysLeuArgGlyThrPheSerAlaAsp 185  
QY 535 AGAATGAGATTTGTTTCTGCAATTCACATCACTATAATATGAGCTGCCGATGAT 476  
DB 186 LysAsnGluLeuLeuPheGlnGluPheGlyIleAsnTyrAsnAsnLeuProAlaMetTyr 205  
QY 475 AGAAAGGGAGCTGTGTTGATATGCGAGAGAGTGGATGAAGTATGACAAAGAAATTAAG 416  
DB 206 ArgLysGlyThrIleLeuLeuArgLysArgVal----- 216  
QY 415 CTGCCAAGAAATGGAAGGAAAAAGATGGCAGTACCCGGACAGGACAAAGCCAGTG 356  
DB 217 -----IleLeuGlyGluLys-----SerArgGlnAlaValVal 227  
QY 355 CCCTTGCACTGCGATATCATCGGGATGCTTCTGGAAGGAACATCCAGAGATTCTA 299  
DB 228 ProLeuHisGluAspLeuIleSerSerGlnPheTrpLysGluHisThrGluIleLeu 246

## RESULT 7

AAG39161  
ID AAG39161 standard; Protein; 497 AA.  
XX  
AC AAG39161;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48413.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130049.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144431.  
PR 19-JUL-1999; 99US-0144432.  
PR 19-JUL-1999; 99US-0144433.

PR 19-JUL-1999; 99US-0144433.  
PR 19-JUL-1999; 99US-0144434.  
PR 19-JUL-1999; 99US-0144435.  
PR 20-JUL-1999; 99US-0144435.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.





```
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 03-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 17-JUL-1999; 99US-0144086.
PR 18-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145219.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.

PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1.52e-05
Score: 129.50
Percent Similarity: 51.96%
Best Local Similarity: 31.37%
Query Match: 11.72%
DB: 21
Length: 502
Matches: 32
Conservative: 21
Mismatch: 40
Indels: 9
Gaps: 2

US-09-914-191-1 (1-598) x AAG39160 (1-502)

QY 595 CAATCTGGACTAACACCAAGCCAGGAGATTACAGGAACTCTTNCAGCAGAC 536
Db 148 LysSerGlyLysSerValserGluThrGlnGluLeuLysAspThrGlnLysGlnGln 167
QY 535 AAGAATGAGATTGTTGTTTCTGAATTCACATCAATCAATGATGATGATGATGAT 476
Db 168 LysAsnGluLeuLeuPheGlnLysPheGlyLysAsnTyrLysThrLeuProGluLeuPhe 187
QY 475 AGGAAAGGAGCTGTTGTTGATATGCGAGAGGTGGATGATGATGATGATGATGAT 416
Db 188 ArgGlnGlySerCysLeuPheLysLysValGluGluThrVal 202
QY 415 CTGCCAACAGAAATGGAAGAAAAAGATGGCAGTGACCCGACCCAGGACCAAGCAGTG 356
Db 203 -----LysHisAspGluAsnGlyAsnProValLysArgLeuArgLysalaVal 219
QY 355 CCCTTGCACTGCGAT---ATCATCGGGGATGCTTTCTGGAAGGAACATCCAGAGATTCTA 299
Db 99US-0152363.
```

Db 220 PheValHisSerGluAsnIleAlaGlyArgSerPheTrpAsnGluClnProSerLeuTyr 239  
QY 298 GATGAA 293  
Db : : : : :  
240 AsnAep 241  
RESULT 9  
AAG39159  
ID AAG39159 standard; Protein; 520 AA.  
XX  
AC AAG39159;  
XX  
DT 18-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 48411.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136332.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 02-JUL-1999; 99US-0142154.  
PR 06-JUL-1999; 99US-0142055.  
PR 08-JUL-1999; 99US-0142390.  
PR 09-JUL-1999; 99US-0142803.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.



PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139461.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142290.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145513.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 26-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0158293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.

PR 26-OCT-1999; 99US-01611359.  
 PR 26-OCT-1999; 99US-01611360.  
 PR 26-OCT-1999; 99US-01611361.  
 PR 28-OCT-1999; 99US-01611920.  
 PR 28-OCT-1999; 99US-01611992.  
 PR 28-OCT-1999; 99US-01611993.  
 PR 29-OCT-1999; 99US-01612142.  
 PR 29-OCT-1999; 99US-01612142.

## Alignment Scores:

Pred. No.: 2,2e-05 Length: 470  
 Score: 128.00 Matches: 32  
 Percent Similarity: 56.57% Conservatives: 24  
 Best Local Similarity: 32.32% Mismatches: 33  
 Query Match: 11.58% Indels: 10  
 DB: 21 Gaps: 3

## US-09-914-191-1 (1-598) x AAG48957 (1-470)

QY 595 CAATCTGGACTACACCAAGCCAGGAGGATTACAGGAACTCTTNCAGCAGAC 536  
 Db 379 LysSerGlyLysSerLysIleGlnAlaGlnAspTyrLeuLysGlyThrGlnThrArgGlu 398  
 QY 535 AAGAATCAGATTGTGTTTCTGAATCAACATCAACTATAATATGAGCTGCCGATGAT 476  
 Db 399 LysAsnGluLeuLeuSerGlnGlnPheGlyIleGluTyrAsnSerLeuProValIlePhe 418  
 QY 475 AGAAAGGGACTGTGTGATATGCGCAGAGGTGGATGATGACAAAGAAATTAAG 416  
 Db 419 ArgMetGlySerSerValPhe-----ArgLeuLysGluGlyValThrGluGlu----- 434  
 QY 415 CTGCCACAGAAATGGAAGAAAGATGCGCAGTGCACCGGACGAGCAGACCCAGTG 356  
 Db 435 ---AsnGlyGluValSerGlyLysGlnVal-----GluAlaGluValGly 448  
 QY 355 CCCTTGCACTGCGATATCATCGGGGATGCTTCTGGAAGGAACATCCAGAGATTCTA 299  
 Db 449 ValAspTyrSerAsnIleIleAspGlnCysPheTrpGlnGlnHisProHisIleLeu 467

## RESULT 11

AAG48956  
 ID AAG48956 standard; Protein; 505 AA.

XX AC AAG48956;  
 XX DT 18-OCT-2000 (first entry)  
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 61881.

XX KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

## Arabidopsis thaliana.

XX EP1033405-A2.  
 XX PD 06-SEP-2000.  
 XX PF 25-FEB-2000; 2000EP-0301439.  
 XX PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 23-MAR-1999; 99US-0123548.  
 PR 29-MAR-1999; 99US-0125788.  
 PR 29-MAR-1999; 99US-0126264.  
 PR 01-APR-1999; 99US-0126785.  
 PR 06-APR-1999; 99US-0127462.  
 PR 08-APR-1999; 99US-0128214.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.  
 PR 28-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 30-APR-1999; 99US-0132407.  
 PR 04-MAY-1999; 99US-0132484.  
 PR 05-MAY-1999; 99US-0132485.  
 PR 06-MAY-1999; 99US-0132486.  
 PR 06-MAY-1999; 99US-0132487.  
 PR 07-MAY-1999; 99US-0132863.  
 PR 11-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134768.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137502.  
 PR 07-JUN-1999; 99US-0137724.  
 PR 08-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139492.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.

```
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145114.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147280.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 18-SEP-1999; 99US-0154018.
PR 18-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 2,25e-05
Score: 128.00
Percent Similarity: 56.57%
Best Local Similarity: 32.32%
Query Match: 11.58%
DB: 21

US-09-914-191-1 (1-598) x AAG48956 (1-505)
QY 595 CAATCTGACTAACAACAGCCCAAGGAGATTACAGGAACTCTTNCAGCAGAC 536
Db 414 LysSerGlyLysSerLysLysileGlnAlaGlnAspTyrLeuLysGlyThrGlnThrArgGlu 433
QY 535 AAGAATGAGATTGTTGTTTCTGAATCAACATCAACTAATAATAAGTGCAGTGCAGTAT 476
Db 434 LysAsnGluLeuLeuSerGlnGlnPheGlyileGluTyrAsnSerLeuProValIlePhe 453
QY 475 AGGAAGGAGCTGTGTTGATATGGCAGAGGTGGATGAAGTATGATGATGACAAAGAAATAAG 416
Db 454 ArgMetGlySerSerValPhe-----ArgLeuLysGluGlyValThrGluGlu----- 469
QY 415 CTGCCAACAGAAATGGAAGGAAAAAGATGGCAGTGACCCGACGAGCAAGCAAGCCAGTG 356
Db 470 ---AsnGlyGluValSerGlyLysGlnVal-----GluAlaGluValGly 483
QY 355 CCCTTGCACTGCGATATCATCGGGATGCTTTCTGGAAGGAACATCCAGAGATTCTTA 299
Db 484 ValAspTyrSerAsnIleIleAspGlnCysPheTyrGlnGlnHisProHisIleLeu 502

RESULT 12
AAG48955
ID AAG48955 standard; Protein; 523 AA.
XX
AC AAG48955;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 61880.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
```

XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140395.  
PR 28-JUN-1999; 99US-0140823.  
PR 28-JUN-1999; 99US-0140931.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142330.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.



```

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161932.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 2,27e-05
Score: 128.00
Length: 523
Matches: 32
Percent Similarity: 56.57%
Conservative: 24
Best Local Similarity: 32.32%
Mismatch: 33
Query Match: 11.58%
Indels: 10
Gaps: 3
DB: 21

US-09-914-191-1 (1-598) x AAG48955 (1-523)

QY 595 CAATCTGGACTACACAGGAGGATACAGGAACTCTTCAGGAGAC 536
DB 432 LysSerGlyLysSerLysIleGlnAlaGlnAspTyrLeuLysGlyThrGlnThrArgGlu 451
QY 535 AAGAATGAGATTGTTGTTCTGAATCAACATCAACTATAATATAGCTGCCGATGTAT 476
DB 452 LysAsnGluLeuSerGlnGlnPheGlyIleGluTyrAsnSerLeuProValIlePhe 471
QY 475 AGGAAAGGAGCTGTGTTGATATGCGAGAGGTGGATGATGATGACAAAGAAATTAAG 416
DB 472 ArgMetGlySerValPhe-----ArgLeuLysGluGlyValThrGluGlu----- 487
QY 415 CTGCCACAGAAATGAGAAAGAAAGATGCGAGTGCACCGGACAGGACAAAGCCAGTG 356
DB 488 ---AsnGlyGluValSerGlyLysGlnVal-----GluAlaGluValGly 501
QY 355 CCCTTGACGCGCATATCATCGGGATGCTTTCTGGAAGGAAACATCCAGAGATTCTA 299
DB 502 ValAspTyrSerAsnIleIleAspGlnCysPheTrpGlnGlnHisProHisIleLeu 520

RESULT 13
ID ABB92421
AC ABB92421;
XX
XX
XX 31-MAY-2002 (first entry)

```

```

XX Herbicidally active polypeptide SEQ ID NO 1632.
DE DE
XX Herbicidal; plant; agriculture; herbicide.
KW KW
XX Arabidopsis thaliana.
OS OS
XX WO200210210-A2.
PN PN
XX 07-FEB-2002.
PD PD
XX 28-AUG-2001; 2001WO-EP09892.
PF PF
XX 28-AUG-2001; 2001WO-EP09892.
PR PR
XX (FARB ) BAYER AG.
PA PA
XX Tietjen K, Weidler M;
PI PI
XX WPI; 2002-269010/31.
DR DR
XX
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
PT
XX Claim 5; SEQ ID NO 1632; 261pp + Sequence Listing; English.
PS PS
XX The invention relates to identifying target proteins
XX (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
XX SQ Sequence 694 AA;

Alignment Scores:
Pred. No.: 0.386
Score: 90.50
Length: 694
Percent Similarity: 38.89%
Conservative: 33
Best Local Similarity: 30.56%
Mismatch: 46
Query Match: 8.26%
Indels: 20
DB: 23
Gaps: 3

US-09-914-191-1 (1-598) x ABB92421 (1-694)

QY 74 CCAACACCCGCAACAGGTGTTCTGCTTAAACAGAGTAAAGATACACACCCCATCCATC 133
DB 160 ProProProProSerProAlaThrProSerThrProProArgSerProProPro----- 177
QY 134 CTTCTCTCCCTGTTCCCTCCCACTTGAGTTGTTCTATTCACACAGTGCTCTGGTG 193
DB 178 ProSerThrProThrProProProArg----- 186
QY 194 GTAGGATGCTACAGCACCTAAGCAAGGAGCCCTGGAGGTGGAGGGCTTGCATGGT 253
DB 187 ValGlySerLeuSerProProProProAlaSerPro-SerGlyGlyArgSerProSerTh 206
QY 254 TAAGCACACAGCAACTGAAGCGCAAAAGGTCAGCTGTCTTCACTTAGAATCTCTGGATG 313
DB 206 rProSerThrThrProGlySerSerProProAlaGlnSerSerLysGluLeuSer----- 224
QY 314 TTCTCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGAGGCAAGGCTTCTCTGGT 373
DB 225 -----LysGlyAlaMetValGlyIleAlaIleGlyGlyPheValLeuLe 240
QY 374 CCGGTCACTGCCATCTTTT 395
DB 240 uValAlaLeuAlaLeuIlePhe 247

```

RESULT 14  
AAG58628 standard; Protein; 128 AA.  
XX AC AAG58628;  
DT 18-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 75712.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134321.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 18-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138340.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145216.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.



SQ Sequence 982 AA;

## Alignment Scores:

Pred. No.:	4.33	Length:	982
Score:	81.50	Matches:	33
Percent Similarity:	39.81%	Conservative:	10
Best Local Similarity:	30.56%	Mismatches:	38
Query Match:	7.38%	Indels:	27
DB:	19	Gaps:	5

US-09-914-191-1 (1-598) x AAW77288 (1-982)

QY	389	GATGGCAGTCAACCGGACAGGCAAGCCAGTGCCTTCACCTGCATATCATCGG---	333
DB	699	AspGluSerAspGluSerLeuAspGluSerProHisArgAspGluArgPro	718
QY	332	---GGATGCTTTCTGGAAGGAACA-----TCCAGA	306
DB	719	IleSerCysThrProGlySerAsnSerLeuGlnLeuSerProAlaSerLeuSerArg	738
QY	305	GATTCTAGATGAACACAGCTGACCCCTTTTCGGCTTCAGTTCTCGT-----	261
DB	739	AspGlyArgAspLeuValLysAspLysGlnArgPheValProAsnLeuValAsnAsnGlu	758
QY	260	-----GTGCTTAACCATGCAAGCCCTCCACCTCCAGGGCTCCTTGCCTTAG	213
DB	759	ThrTyrGlyThrIleIleAsnThrSerSerProValSerLeuSerSerSerAlaPro---	777
QY	212	GTGGCTGTAGCATCCCTTACCACCGACACACTGGTGGCAATGACAACTCAAAGTTGGGA	153
DB	778	-----ProLeuProProArgAsnLeuValGlnProSerAlaLeu---AlaGly	792
QY	152	GGGGAACAGGAAGGAGGATCG	129
DB	793	LeuThrGlnGlySerProGlyTrp	800

Search completed: July 9, 2003, 12:48:12  
Job time : 44.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 9, 2003, 12:46:48 ; Search time 13.5 Seconds  
(without alignments)

2606.653 Million cell updates/sec

Title: US-09-914-191-1

Perfect score: 1096

Sequence: 1 tgggaatagttctgttctta.....gggttagtcacagattgttg 598

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/US09914191/runat\_09072003\_112350\_12911/app.query.fasta\_1.775  
-DB=Issued Patents AA -GPM=fastan -SUFFIX=ra1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09914191 @CIGN 1.1 28 @runat 09072003\_112350\_12911 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	6.9	419	4	US-09-461-697-389
2	76	6.9	488	4	US-09-461-697-375
3	76	6.9	490	4	US-09-461-697-373
4	76	6.9	490	4	US-09-056-285A-10
5	76	6.9	1167	4	US-09-008-097-6
6	75.5	6.8	393	1	US-08-429-742-2
7	72.5	6.6	453	2	US-08-484-126-3
8	72.5	6.6	1012	4	US-08-811-481-16
9	72	6.6	375	4	US-09-461-697-395
10	71	6.4	247	2	US-08-463-311-2
11	71	6.4	247	4	US-09-140-804-8
12	71	6.4	247	4	US-09-118-408-3

Sequence 3, Appli	247	6.4	4	US-09-506-855-3
Sequence 43, Appli	331	6.4	4	US-09-108-020-43
Sequence 2, Appli	2201	6.4	4	US-08-952-981A-2
Sequence 32, Appli	2620	6.4	1	US-08-324-977-32
Sequence 32, Appli	2620	6.4	2	US-08-384-616-32
Sequence 32, Appli	2620	6.4	2	US-08-904-686A-32
Sequence 32, Appli	2620	6.4	4	US-09-315-850-32
Sequence 36, Appli	2621	6.4	1	US-08-324-977-36
Sequence 36, Appli	2621	6.4	2	US-08-384-616-36
Sequence 36, Appli	2621	6.4	2	US-08-904-686A-36
Sequence 36, Appli	2621	6.4	4	US-09-315-850-36
Sequence 2, Appli	3010	6.4	1	US-08-324-977-2
Sequence 14, Appli	3010	6.4	1	US-08-324-977-14
Sequence 2, Appli	3010	6.4	2	US-08-384-616-2
Sequence 14, Appli	3010	6.4	2	US-08-904-686A-14
Sequence 2, Appli	3010	6.4	4	US-09-315-850-2
Sequence 14, Appli	3010	6.4	4	US-09-315-850-14
Sequence 2, Appli	3010	6.4	4	US-09-518-914-2
Sequence 2, Appli	2431	6.4	1	US-07-920-281C-2
Sequence 2, Appli	2431	6.4	4	US-08-466-277-2
Sequence 2, Appli	3010	6.3	4	US-09-014-416-3
Sequence 2, Appli	283	6.3	4	US-09-222-939-2
Sequence 358, App	274	6.3	4	US-09-149-476-358
Sequence 23, Appli	387	6.3	3	US-08-689-421-23
Sequence 23, Appli	387	6.3	4	US-09-389-528-23
Sequence 23, Appli	387	6.3	4	US-09-181-827A-23
Sequence 29, Appli	516	6.3	3	US-08-689-421-29
Sequence 29, Appli	516	6.3	4	US-09-389-528-29
Sequence 29, Appli	516	6.3	4	US-09-181-827A-29
Sequence 6, Appli	638	6.3	4	US-09-376-781-6
Sequence 29, Appli	2291	6.2	2	US-08-286-819A-29

ALIGNMENTS

RESULT 1

US-09-461-697-389  
; Sequence 389, Application US/09461697  
; Patent No. 6277974  
; GENERAL INFORMATION:  
; APPLICANT: COGENT NEUROSCIENCE, Inc.  
; APPLICANT: Lo, Donald C.  
; APPLICANT: Barney, Shawn  
; APPLICANT: Thomas, Mary Beth  
; APPLICANT: Portbury, Stuart D.  
; APPLICANT: Puranam, Kasturi  
; APPLICANT: Katz, Lawrence C.  
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING  
; TITLE OF INVENTION: CELL DEATH  
; FILE REFERENCE: 10001-005-999  
; CURRENT APPLICATION NUMBER: US/09/461,697  
; CURRENT FILING DATE: 1999-12-14  
; NUMBER OF SEQ ID NOS: 466  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 389  
; LENGTH: 419  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-461-697-389

Alignment Scores:  
Pred. No.: 1.71  
Score: 76.00  
Percent Similarity: 33.70%  
Best Local Similarity: 26.09%  
Query Match: 6.93%  
DB: 4  
Length: 419  
Matches: 24  
Conservative: 7  
Mismatch: 21  
Indels: 40  
Gaps: 4

US-09-914-191-1 (1-598) x US-09-461-697-389 (1-419)

```
QY 323 GAAGCATCCCGATGATATCGCAGTCAAGGCACTGGCTTTGTCTCT----- 370
Db 39 GluAlaAspArgAspIleMetThrLeuAlaAsnHisTrpAsnCysProValLeuSerSer 58
QY 371 -----GGTCCGGGTCACTGCCATCTTTTTCCTTCCAT 403
Db 59 AspSerAspPheCysIlePheAspLeuLysThrGlyPheCysProLeuAsnSerPheGln 78
QY 404 TTC-----TGTGGCAG 415
Db 79 TrpArgAsnMetAsnThrIleLysGlyThrGlnAsnTyrIleProAlaLysCysPheSer 98
QY 416 CTTAAATTTCTTTGTCACTCATCCAC-----CTTCTGCCATAT----- 457
Db 99 LeuAspAlaPheCysHisHisPheSerAsnMetAsnLysAlaLeuLeuProLeuPheAla 118
QY 458 -----CAACACAGTCCCTTTTCTTATA 478
Db 119 ValLeuCysGlyAsnAspHisValAsnLeuProIle 130

RESULT 2
US-09-461-697-375
; Sequence 375, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 375
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-375

Alignment Scores:
Pred. No.: 1.82 Length: 488
Score: 76.00 Matches: 24
Percent Similarity: 33.70% Conservative: 7
Best Local Similarity: 26.09% Mismatches: 21
Query Match: 6.93% Indels: 40
DB: 4 Gaps: 4

US-09-914-191-1 (1-598) x US-09-461-697-375 (1-488)
QY 323 GAAGCATCCCGATGATATCGCAGTCAAGGCACTGGCTTTGTCTCT----- 370
Db 108 GluAlaAspArgAspIleMetThrLeuAlaAsnHisTrpAsnCysProValLeuSerSer 127
QY 371 -----GGTCCGGGTCACTGCCATCTTTTTCCTTCCAT 403
Db 128 AspSerAspPheCysIlePheAspLeuLysThrGlyPheCysProLeuAsnSerPheGln 147
QY 404 TTC-----TGTGGCAG 415
Db 148 TrpArgAsnMetAsnThrIleLysGlyThrGlnAsnTyrIleProAlaLysCysPheSer 167
QY 416 CTTAAATTTCTTTGTCACTCATCCAC-----CTTCTGCCATAT----- 457
Db 168 LeuAspAlaPheCysHisHisPheSerAsnMetAsnLysAlaLeuLeuProLeuPheAla 187
QY 458 -----CAACACAGTCCCTTTTCTTATA 478
```

```
Db 188 ValLeuCysGlyAsnAspHisValAsnLeuProIle 199

RESULT 3
US-09-461-697-373
; Sequence 373, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 373
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-373

Alignment Scores:
Pred. No.: 1.82 Length: 490
Score: 76.00 Matches: 24
Percent Similarity: 33.70% Conservative: 7
Best Local Similarity: 26.09% Mismatches: 21
Query Match: 6.93% Indels: 40
DB: 4 Gaps: 4

US-09-914-191-1 (1-598) x US-09-461-697-373 (1-490)
QY 323 GAAGCATCCCGATGATATCGCAGTCAAGGCACTGGCTTTGTCTCT----- 370
Db 110 GluAlaAspArgAspIleMetThrLeuAlaAsnHisTrpAsnCysProValLeuSerSer 129
QY 371 -----GGTCCGGGTCACTGCCATCTTTTTCCTTCCAT 403
Db 130 AspSerAspPheCysIlePheAspLeuLysThrGlyPheCysProLeuAsnSerPheGln 149
QY 404 TTC-----TGTGGCAG 415
Db 150 TrpArgAsnMetAsnThrIleLysGlyThrGlnAsnTyrIleProAlaLysCysPheSer 169
QY 416 CTTAAATTTCTTTGTCACTCATCCAC-----CTTCTGCCATAT----- 457
Db 170 LeuAspAlaPheCysHisHisPheSerAsnMetAsnLysAlaLeuLeuProLeuPheAla 189
QY 458 -----CAACACAGTCCCTTTTCTTATA 478
Db 190 ValLeuCysGlyAsnAspHisValAsnLeuProIle 201

RESULT 4
US-09-056-285A-10
; Sequence 10, Application US/09056285A
; Patent No. 6403307
; GENERAL INFORMATION:
; APPLICANT: Stone, Edwin M.
; APPLICANT: Sheffield, Val C.
; APPLICANT: Alward, Wallace L.M.
; APPLICANT: Fingert, John
; TITLE OF INVENTION: GLAUCOMA THERAPEUTICS AND DIAGNOSTICS
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
```







```

; FEATURE:
; NAME/KEY: xenotropic gp70 protein
; US-08-484-126-3

Alignment Scores:
Pred. No.: 4,6 Length: 453
Score: 72.50 Matches: 44
Percent Similarity: 33.85% Conservative: 21
Best Local Similarity: 22.92% Mismatches: 59
Query Match: 6.61% Indels: 68
DB: 2 Gaps: 11

US-09-914-191-1 (1-598) x US-08-484-126-3 (1-453)
QY 113 GATACACACCCCATCC-----ATCCCTTCTTCCCTGTTCCCTC 154
DB 255 AspGlnLeuProSerGlnProValGlnLeuProAArgProHisProPro 274
QY 155 CCACTTGAGTTGTGTCATTCGCACAGTGTCTGGTGGTGGTAGGATGCTACAGCCACT 214
DB 275 ProSerGlyThrValSerMetValProGlyAla-----ProPro 287
QY 215 AAGCAGAGGAGCCCTGGAGGTGGAGGCTTGCATGTTAAGCACACCACTGAAGC 274
DB 288 ProSerGlnProGlyThrGlyAspArgLeuLeuAsn----- 300
QY 275 GCAAAAGGTCAGCTGTCTTCATCTAGATCTCTGGATGTTCTCCAGAAAGCATCCC 334
DB 301 -----LeuValGluGlyAlaTyrglnAlaLeuAsnLeuThrSerPro 314
QY 335 GATGATATCGAGTGCAGGACATCGCTTGT----- 367
DB 315 Asp-----LysThrGlnGluCysTrpLeuCysLeuValSerGlyProProTyrglu 332
QY 368 -----CCTGGTCGGGTCACTGCCACTT 391
DB 333 GlyValAlaValLeuGlyThrTyrglnHisThrSerAlaProAlaAsnCysVal 352
QY 392 TTTTCTTCCATTTCTGTGGCAGCTTAATTTCTTTGTGTCATCTCATCCACTTCTG 451
DB 353 AlaSerGlnHis-----LysLeuThrLeu-SerGluValThrGlyGlnGlyLeuCy 369
QY 452 CCATATCAACACAGCTCCCTTCCATATACCGGAGCTCATATTATATGTTGATGTTGAA 511
DB 369 s---ValGlyAlaValProLysThrHisGlnAla-----LeuCysAs 382
QY 512 TTCAGAAAACAAATCTCA-----TTCTTGTCTGCTGNAAGAGTTTCCCTGTAA 559
DB 382 nThrThrGlnLysThrSerAspGlySerTyrglnLeuAlaAla-----ProAlaG 399
QY 560 TCTCCCTTGGCTGTGT---ACTGGTGTAGTCCA 590
DB 399 yThrIleTrpAlaCysAsnThrGlyLeuThrPro 410

RESULT 8
US-08-811-481-16
; Sequence 16, Application US/08811481
; Patent No. 6300093
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; APPLICANT: Jelinek, Laura J.
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Hagopian, William A.
; APPLICANT: LaGasse, James M.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,481
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-811-481-16

Alignment Scores:
Pred. No.: 6,39 Length: 1012
Score: 72.50 Matches: 26
Percent Similarity: 37.37% Conservative: 11
Best Local Similarity: 26.26% Mismatches: 50
Query Match: 6.61% Indels: 12
DB: 4 Gaps: 3

US-09-914-191-1 (1-598) x US-08-811-481-16 (1-1012)
QY 129 CCATCCCTTCCCTTCCCTGTTCCCTCCCACTTGAGTTGTGTCATTCGCACCACTGTCCT 188
DB 18 ProAlaAlaProSerSerValPro-----HisGlyArgGlnLeuPro 31
QY 189 GGTGTGTAGGTGTGTACAGCCACCTAAGCGAAGAGCCCTGGAGGTGGAGGGTGTGC 248
DB 32 GlyArgLeuGlyCysLeu-----LeuGluGluGlyLeuCysGlyAlaSerGluAlaCys 49
QY 249 ATGGTTTAAAGCACACCAAGAACTGAAGCGCAAAAG-----GGTCAGCTGTCTTCATCT 299
DB 50 ValAsnAspGlyValPheGlyArgCysGlnLysValProAlaMetAspPheTyrgly 69
QY 300 AGAATCTCTGGATGTTCTTCAGAAAGCATCCCGATGATATCGCAGTGCAGAGGCACT 359
DB 70 GluValSerProValAlaLeuGlnArgLeuArgValAlaLeuGlnLysLeuSerGlyThr 89
QY 360 GCTTTGTCTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 414
DB 90 GlyPhe-ThrTrpGlnAspTyrglyThrGlnTyrglyValMetAspGlnGluLeuAla 107

RESULT 9
US-09-461-697-395
; Sequence 395, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENET NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING

```

```
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461.697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 395
; LENGTH: 375
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-395

Alignment Scores:
Pred. No.: 4.99 Length: 375
Score: 72.00 Matches: 22
Percent Similarity: 32.56% Conservative: 6
Best Local Similarity: 25.58% Mismatches: 18
Query Match: 6.57% Indels: 40
DB: 4 Gaps: 4

US-09-914-191-1 (1-598) x US-09-461-697-395 (1-375)
QY 341 ATCGCAGTCAAGGCGCACTGCTTGTCTCT----- 370
Db 1 MetThrLeuAlaAsnHisTrpAsnCysProValLeuSerSerAspSerAspPheCysIle 20
QY 371 -----GGTCGGGTCACTGCATCTTTTTCCTTCCATTTTC----- 406
Db 21 PheAspLeuIysThrGlyPheCysProLeuAsnSerPheGlnTrpArgAsnMetAsnThr 40
QY 407 -----TGTTGGCAGCTTAATTTCTTTTCTTCAT 433
Db 41 IleIysGlyThrGlnAsnTyrlleProAlaIysCysPheSerLeuAspAlaPheCysHis 60
QY 434 CACTTCATCCAC-----CTTCTGCCATAT-----CAA 460
Db 61 HisPheSerAsnMetAsnLysAlaLeuLeuProLeuPheAlaValLeuCysGlyAsnAsp 80
QY 461 CACAGTCCCTTTCTCTATA 478
Db 81 HisValAsnLeuProIle 86

RESULT 10
US-08-463-911-2
; Sequence 2, Application US/08463911
; Patent No. 5869330
; GENERAL INFORMATION:
; APPLICANT: Scherer, Philipp E.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
; TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,911
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI95-05
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-911-2

Alignment Scores:
Pred. No.: 5.42 Length: 247
Score: 71.00 Matches: 15
Percent Similarity: 51.28% Conservative: 5
Best Local Similarity: 38.46% Mismatches: 11
Query Match: 6.43% Indels: 8
DB: 2 Gaps: 2

US-09-914-191-1 (1-598) x US-08-463-911-2 (1-247)
QY 241 CTCCACCTCCAGGGCTCCTTGCCTTAGTGGCTGTAGCAT---CCCTACCACCCAGGA 185
Db 32 ValProProIysGlyThrCysAlaGlyTrpMetAlaGlyIleProGlyHisProGly 51
QY 184 CACTGTGCGAATGACACAACTCAAGTTGGAGGGGAACAGGGAAGGAGGATGA 128
Db 52 His-----AsnGlyThrProGlyArgAspGlyArgAspGly 63

RESULT 11
US-09-140-804-8
; Sequence 8, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49
; CURRENT APPLICATION NUMBER: US/09/140,804
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056,983
; EARLIER FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-140-804-8

Alignment Scores:
Pred. No.: 5.42 Length: 247
Score: 71.00 Matches: 15
Percent Similarity: 51.28% Conservative: 5
Best Local Similarity: 38.46% Mismatches: 11
Query Match: 6.43% Indels: 8
DB: 4 Gaps: 2

US-09-914-191-1 (1-598) x US-09-140-804-8 (1-247)
QY 241 CTCCACCTCCAGGGCTCCTTGCCTTAGTGGCTGTAGCAT---CCCTACCACCCAGGA 185
Db 32 ValProProIysGlyThrCysAlaGlyTrpMetAlaGlyIleProGlyHisProGly 51
QY 184 CACTGTGCGAATGACACAACTCAAGTTGGAGGGGAACAGGGAAGGAGGATGA 128
Db 52 His-----AsnGlyThrProGlyArgAspGlyArgAspGly 63

RESULT 12
US-09-118-408-3
; Sequence 3, Application US/09118408A
; Patent No. 6265544
; GENERAL INFORMATION:
```



```
; Sequence 2, Application US/08952981A
; Patent No. 6383768
; GENERAL INFORMATION:
; APPLICANT: DE FRANCESCO, Raffaele
; APPLICANT: TOMEI, Licia
; APPLICANT: BEHRENS, Sven-Erik
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA
; TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE
; TITLE OF INVENTION: ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)
; FILE REFERENCE: IT00002P
; CURRENT APPLICATION NUMBER: US/08/952,981A
; CURRENT FILING DATE: 1998-03-23
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: cdna clone pCD (38-9.4)
US-08-952-981A-2
```

```
Alignment Scores:
Pred. No.: 15.2 Length: 2201
Score: 70.50 Matches: 24
Percent Similarity: 41.57% Conservative: 13
Best Local Similarity: 26.97% Mismatches: 33
Query Match: 6.38% Indels: 19
DB: 4 Gaps: 4
```

US-09-914-191-1 (1-598) x US-08-952-981A-2 (1-2201)

```
QY 450 AGAAGTGGATGAAGTCATGACAAAGAAATTA----- 418
Db 1183 LysThrTrpLeu-----GlnSerLysLeuLeuProGlnLeuProGlyValProphe 1199

QY 417 ---AGTCGCCACAGAAATGGAAGAAAGAAAGATGCGCAGTGACCCGGACCCAGGACAAAGC 361
Db 1200 PheSerCysGlnArgGlyTyrLysGlyVal---TipArgGlyAspGlyIleMetGlnThr 1218

QY 360 CAGTGCCTTGCCTGCGATATCATCGGGATGCTTCTCGAAGGAACATCCAGAGATTC 301
Db 1219 ThrCysProCysGlyAlaGlnIleThrGlyHisValLysAsnGlySerMetArgIleVal 1238

QY 300 TAGATGAAGACAGCTGAC-----CCTTTTCGCTTCAGTTCTCGTGTGCT 256
Db 1239 GlyProLysThrCysSerAsnThrTrpHisGlyThrPheProIleAsnAlaTyrThrThr 1258

QY 255 TAACCATGCAAGCCCTCCACCTCCCA 229
Db 1259 GlyProCysThrProSerProAlaPro 1267
```

Search completed: July 9, 2003, 12:51:58  
Job time : 19.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 9, 2003, 12:48:19 ; Search time 31 Seconds

(without alignments)

4491.550 Million cell updates/sec

Title: US-09-914-191-1

Perfect score: 1096

Sequence: 1 ttggaatagttcttcttctta.....gggttagtcagattgttg 598

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 445758 seqs, 116419773 residues

Total number of hits satisfying chosen parameters: 891516

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool/US09914191/runat\_09072003\_112351\_12946/app.query.fasta\_1.775  
-DB=Published Applications\_AA -QFMT=fastan -SURFIX=rapb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09914191.cgn 1 1 17 @runat\_09072003\_112351\_12946  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA:

1: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
2: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
6: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	455	41.2	85	9	US-10-106-698-5245
2	88.5	8.1	457	9	US-09-934-455-302
3	80	7.3	233	10	US-09-925-300-1529
C 4	78	7.1	833	10	US-09-844-281-1

Alignment Scores: 4.43e-40 Length: 85  
Pred. No.: 455.00 Matches: 85  
Score: 100.00%  
Percent Similarity: 100.00%  
Conservative: 0

#### ALIGNMENTS

RESULT 1

US-10-106-698-5245

; Sequence 5245, Application US/10106698

; Publication No. US20030109690A1

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides

; FILE REFERENCE: PA005PI

; CURRENT APPLICATION NUMBER: US/10/106,698

; PRIOR FILING DATE: 2002-03-27

; PRIOR APPLICATION NUMBER: PCT/US00/26524

; PRIOR FILING DATE: 2000-09-28

; PRIOR APPLICATION NUMBER: US 60/157,137

; PRIOR FILING DATE: 1999-09-29

; PRIOR APPLICATION NUMBER: US 60/163,280

; PRIOR FILING DATE: 1999-11-03

; NUMBER OF SEQ ID NOS: 8564

; SOFTWARE: PatentIn Ver. 3.0

; SEQ ID NO 5245

; LENGTH: 85

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-106-698-5245

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.18% Indels: 0
DB: 9 Gaps: 0

US-09-914-191-1 (1-598) x US-10-106-698-5245 (1-85)

QY 541 GCACAGAAGTATGATTTTCTTCTGAATCAACATCACTATATATATAGCTGCG 482
Db 1 AlaAspLysAsnGluLeuLeuPheSerGluPheAsnIleAsnIleAsnGluLeuPro 20
QY 481 ATGTATAGGAAAGCGACTGTCTGATATGCGAAGGTGGATGAAGTGCATCAACAAAGAA 422
Db 21 MetTyArgLysGlyThrValLeuLeuLeuLeuLeuValAspGluValMetThrLysGlu 40
QY 421 ATTAAGCTGCCAACAGAAATGGAAGGAAAAAGATGCCAGTGCACCGGACAGACAAAG 362
Db 41 IleLysLeuProThrGluMetGluGlyLysMetAlaValThrArgThrArgThrLys 60
QY 361 CCAGTGCCTTGCACCTCGATATCATCGGGGATGCTTCTTGGAGGAACATCCAGAGATT 302
Db 61 ProValProLeuHisCysAspIleIleGlyAspAlaPheTrpLysGluHisProGluIle 80
QY 301 CTAGATGAACAGAC 287
Db 81 LeuAspGluAspSer 85

RESULT 2
US-09-934-455-302
; Sequence 302, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
; APPLICANT: Adam, Luc
; APPLICANT: Creelman, Robert
; APPLICANT: Dubell, Arnold
; APPLICANT: Heard, Jacqueline
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Keddle, James
; APPLICANT: Pilgrim, Mareha
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Reuber, Lynne
; APPLICANT: Riethmann, Jose Luis
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: Genes for Modifying Plant Traits IV
; FILE REFERENCE: MBI-0025
; CURRENT APPLICATION NUMBER: US/09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/227439
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: MBI-0022
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: MBI-0023
; PRIOR FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 516
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 302
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-934-455-302

Alignment Scores:
Pred. No.: 1.06 Length: 457
Score: 88.50 Matches: 28
Percent Similarity: 35.87% Conservative: 5
Best Local Similarity: 30.43% Mismatches: 32
Query Match: 8.07% Indels: 27
DB: 9 Gaps: 3

US-09-914-191-1 (1-598) x US-09-934-455-302 (1-457)

QY 71 GTTCCACACCGCAAGAGTGTCTGTTAAACAGAGTAAGATACACACCCCATCC 130

```

```

Db 267 ValProLeuAsnGlnThrAsnGluGluAlaGlyThrVal-----SerProLeuProLys 284
QY 131 ATCCCTTCCTTCCTGTTCCCTCCCACTTCACTGAGTTGTGTCTATCGCACAGTCTCTGG 190
Db 285 ValProCysPheProGlyProProThrTrpProTyAlaTrpAsnGlyValSerTrp 304
QY 191 GTGTAGGAGTCTACAGCCACCTAAGCAAGG-----AGCCCTGGG 232
Db 305 ThrIleLeuProPheTyTrpProProAlaIleTyTrpSerCysProGlyValSerProGly 324
QY 233 AGGTGGGGGCTTGCATGTTAAGCACACCAAGAACTGAAGCGCAAAAGGGTCAGCTGTC 292
Db 325 AlaTrpAsnSerPhe----- 329
QY 293 TTCATCTAGAACTCTGGATGTTCTTCCAGAAAGC 328
Db 330 -----ThrTrpMetProGlnProAsnSer 337

RESULT 3
US-09-925-300-1529
; Sequence 1529, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1529
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1529

Alignment Scores:
Pred. No.: 7.12 Length: 233
Score: 80.00 Matches: 27
Percent Similarity: 40.40% Conservative: 13
Best Local Similarity: 27.27% Mismatches: 26
Query Match: 7.30% Indels: 33
DB: 10 Gaps: 2

US-09-914-191-1 (1-598) x US-09-925-300-1529 (1-233)

QY 87 CAAGTGTTCTGCTTAAACAGAGTAAGATACACACCCCATCCATCCCTTCCTTCCCTG 146
Db 85 GlnGlySerLeuValAsnGlnAsnLeuLeuHisHisGlnHisGlnThr----- 100
QY 147 TTCCCTCCCAACTTGAGTTGTCTATCGCACCAAGTGTCTTGGTGGTAGGATGCTAC 206
Db 100 ----- 100
QY 207 AGCCACCTAAGCAAGGAGCCCTGGAGGTGGAGGCTTGCATGTTAAGCACACACAGA 266
Db 101 -----GlnGlyAlaLeuGlyGlySerArgAlaLeu-----SerAsn 112
QY 267 ACTGAAGCGCAAAAGGTCAGTCTTTCATCTAGATCTCTGGATGTTCTTCCAGAAA 326
Db 113 SerValSerAsnMetGlyLeuSerGluSerSerLeuGlySerAlaLysHisGlnGln 132
QY 327 GCATCCCGATGATATCGCAGTGCAGGCAAGGCACTGGCTTTGCTCTGTCTCCGGTCA 381
Db 133 GlnSerProVal-SerGlnSerMetGlnThrLeuSerAspSerLeuSerGlySer 150

RESULT 4
US-09-844-281-1

```

## RESULT 5





US-09-922-261-373

Alignment Scores:  
Pred. No.: 23.5 Length: 490  
Score: 76.00 Matches: 24  
Percent Similarity: 33.70% Conservative: 7  
Best Local Similarity: 26.09% Mismatches: 21  
Query Match: 6.93% Indels: 40  
DB: 10 Gaps: 4

US-09-914-191-1 (1-598) x US-09-922-261-373 (1-490)

QY 323 GAAGCATCCCGATGATGCGAGTGAAGGCGACTGGTTGTCTCT----- 370  
Db 110 GluAlaAspArgAspIleMetThrLeuAlaAsnHisTrpAsnCysProValLeuSerSer 129  
QY 371 -----GGTCCGGTCACTGCCATCTTTTCTCTCCAT 403  
Db 130 AspSerAspPheCysIlePheAspLeuLysThrGlyPheCysProLeuAsnSerPheGln 149  
QY 404 TTC-----TGTTGGCAG 415  
Db 150 TrpArgAsnMetAsnThrIleLysGlyThrGlnAsnTyrlleProAlaLysCysPheSer 169  
QY 416 CTTAATTCTTTGTCTATCTCTCATCCAC-----CTTCTGCCATAT----- 457  
Db 170 LeuAspAlaPheCysHisHisPheSerAsnMetAsnLysAlaLeuLeuProLeuPheAla 189  
QY 458 -----CAACACAGTCCCTTTCCTATA 478  
Db 190 ValLeuCysGlyAsnAspHisValAsnLeuProIle 201

RESULT 9

US-09-750-240-6  
; Sequence 6, Application US/09750240  
; Patent No. US20020103147A1  
; GENERAL INFORMATION:  
; APPLICANT: Hammon, H. K.  
; APPLICANT: Insel, P. A.  
; APPLICANT: Ping, P.  
; APPLICANT: Post, S. R.  
; APPLICANT: Gao, M.

; TITLE OF INVENTION: GENE THERAPY FOR CONGESTIVE HEART

; FILE REFERENCE: FAILURE

; CURRENT APPLICATION NUMBER: US/09/750,240

; PRIOR FILING DATE: 2001-10-12

; PRIOR FILING DATE: 1999-12-27

; PRIOR FILING DATE: 1997-06-16

; PRIOR FILING DATE: 1998-01-16

; PRIOR FILING DATE: 1997-09-05

; PRIOR FILING DATE: 1997-06-16

; PRIOR FILING DATE: 1996-09-05

; NUMBER OF SEQ ID NOS: 13

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 1167

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-750-240-6

Alignment Scores:

Pred. No.: 30.1 Length: 1167  
Score: 76.00 Matches: 32  
Percent Similarity: 41.54% Conservative: 22  
Best Local Similarity: 24.62% Mismatches: 46  
Query Match: 6.93% Indels: 30  
DB: 10 Gaps: 7

US-09-914-191-1 (1-598) x US-09-750-240-6 (1-1167)

QY 132 TCCTTCTCCTCCTGTTCCCTCCCAACTTGAG----- 164  
Db 718 SerCysGlySerLeuPheProLysAlaLeuGlnArgLeuSerArgSerIleValArgSer 737  
QY 165 TTGTGTTCATCCAGCAGTGTCTCTGGGTGGTAGGATGTACAGCCACCTAAGG----- 218  
Db 738 ArgAlaHisSerThrAlaVal-----GlyIlePheSerValLeuLeuValPhe 753  
QY 219 CAAGGAGCCCTGGGAGGTGGAGGGCTTCATGTTAAGCACACC-----AGAACT 269  
Db 754 ThrSerAlaIleAlaAsnMetPheThrCys-----AsnHisThrProIleArgSerCys 771  
QY 270 GAAGCCGAAAGGGTCAGTCTTCATCTAGAACTCTCTGGATGTCTCTCCAGAAAGCA 329  
Db 772 AlaAlaArgMetLeuAsnLeuThrProAlaAspIleThrAlaCysHisLeuGlnLeu 791  
QY 330 TCCCGCATGATATCG-----CAGTCAAGGCGACTGGCTTGTCTGTGTCGG 377  
Db 792 AsnTySerLeuGlyLeuAspAlaProLeuCysGluGlyThr----- 805  
QY 378 GTCACCTGCCATCTTTTCTCTCCATTTCTGTGGCAGCTTAATTTCTTTTGTCTCACT 437  
Db 806 MetProThrCysSerPheProGluValSerIleGlyAsnMetLeuSerLeuLeuAla 825  
QY 438 TCATCCACCTTCTGCCATATCAACACAGTC 467  
Db 826 SerSerValPheLeuHisIleSerSerIle 835

RESULT 10

US-09-984-130-144  
; Sequence 144, Application US/09984130  
; Publication No. US20030055231A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.

; TITLE OF INVENTION: 12 Human Secreted Proteins

; FILE REFERENCE: PF489P2

; CURRENT APPLICATION NUMBER: US/09/984,130

; PRIOR FILING DATE: 2001-10-29

; PRIOR FILING DATE: 2000-10-30

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 2000-04-19

; PRIOR FILING DATE: 1999-10-27

; PRIOR FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 149

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 144

; LENGTH: 393

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-984-130-144

Alignment Scores:

Pred. No.: 25 Length: 393  
Score: 75.50 Matches: 34  
Percent Similarity: 37.04% Conservative: 16  
Best Local Similarity: 25.19% Mismatches: 46  
Query Match: 6.83% Indels: 39  
DB: 9 Gaps: 7

US-09-914-191-1 (1-598) x US-09-984-130-144 (1-393)

QY 592 TCTGGACTAACA-----CCAGTACAGCCCAAGGAGATTACAGGA 551  
Db 54 SerGlyPheThrIlePheLeuAsnGluTyProAlaLeuLysAsnSerLysTyrGln--- 72  
QY 550 ACTCTTNCAGCAGACAAGAAATGAGATTTTGTCTTCTCAATTCACATCACTATAATAAT 491

```
Db 73 LeuLeuHisHisSerAlaAenGlnLeuSerIleThrValProAenValThrLeuGlnAsp 92
QY 490 GAGTCCGATGTATAGAAAGGACTGTCTTATATATGCGAGAGGTGGATGAGTATG 431
Db 93 Glu-----GlyValTyrLysCysLeuHisTyrSerAspSerValSer 106
QY 430 ACAAAGAAATTAAGCTGCCAACAGAAATGGAAGGAAAAAGATGGCAGTGCACCGGACC 371
Db 107 ThrLysGluVal-----LysValIleValLeuAlaThr 117
QY 370 AGGACAAAGCAGTCCCTTGCACATCGATATATCGGGATGCTTTCTGGAAGGAACAT 311
Db 118 ProPheLysProIle---LeuGluAlaSerValIleArgLysGlnAenGlyGluGluHis 136
QY 310 CCAGAGATTCATAGATGAAGACAGTGCACCTTTTGGCTTCAGTTCCTGGTGTCTTAACC 251
Db 137 -----ValValLeuMetCysSerThr 143
QY 250 ATCAAGCCCTCCACCTCCAGGGCTCCTTGGCTTAGTGGCTG 206
Db 144 MetArgSerLysProProGln-----IleThrTrpLeu 155
```

## RESULT 11

```
US-09-925-300-1578
; Sequence 1578, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1578
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (209)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

```
US-09-925-300-1578
Alignment Scores:
Pred. No.: 28.3 Length: 393
Score: 75.00 Matches: 29
Percent Similarity: 39.47% Conservative: 16
Best Local Similarity: 25.44% Mismatches: 45
Query Match: 6.84% Indels: 24
DB: 10 Gaps: 7
```

## US-09-914-191-1 (1-598) x US-09-925-300-1578 (1-393)

```
QY 58 CCAAGGAACCATGTTCCAAACCGCAACCAAGGTGTTCTGCTTAAACAGAGTAAGATAC 117
Db 292 ProArgArgProCysArgSerIleLysGlyAlaTrpLeuCysCysLeuGlnSerArgThr 311
QY 118 ACCACCCCATCATCC-----TTCTTCCCTGTTCCCTCCCACTTGAGT 165
Db 312 ThrArgLeuThrArgProSerThrThrThrPheLeuSerCysSerProPro-----Ala 329
QY 166 TGTGTCTATCGACCATGTCCTGGTGGTAGGATGCTACACCCACCTAAGCAAGGAG 225
Db 330 AlaProLeuHisProSerThrMetGlySerArgSerProProLeuGln---GlyArgAla 348
QY 226 CCCTGGGAGGTGGAGGGCTTGCATGTTAAGCACACCAAGCAAGTGAAGCGCAAAAGGGTC 285
```

```
Db 349 ProGlnPro-----ArgSerTrpThrGlyThrArgArgArgArgLysArgArgMet 365
QY 286 AGCTGTCTTCATCTAGAAATCTCTGGATGTCTCTCCAGAAAGCATCCCGATGATATCGC 345
Db 366 -----MetLysMetGluLysMet-----ArgArgLysSerPro----- 376
QY 346 AGTCAAGGCGACTGCTTGTCTTGTCTCGGTCCGGGTCACTGCCA 387
Db 377 -----SerAlaSerGlyLysGlySerArgPro 385
```

## RESULT 12

```
US-10-184-644-107
; Sequence 107, Application US/10184644
; Publication No. US2003004930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 107
; LENGTH: 2478
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-184-644-107
```

```
Alignment Scores:
Pred. No.: 69 Length: 2478
Score: 73.50 Matches: 29
Percent Similarity: 35.05% Conservative: 5
Best Local Similarity: 29.90% Mismatches: 48
Query Match: 6.65% Indels: 15
DB: 9 Gaps: 4
```

## US-09-914-191-1 (1-598) x US-10-184-644-107 (1-2478)

```
QY 357 TGCCTTGCATCGCATATCATCGGGATGCTTTCTGGNAGGAACATCCAGAGATTCTAG 298
Db 1125 CysThrCysAlaAlaGlyThrAlaGlyCysThrGlyGly----- 1137
QY 297 ATGAAGACAGCTGACCCCTTTTCGGCTTCAGTTCTGGTGTCTTAACCATGCAAGCCCTCC 238
Db 1138 ---GlyAlaCysThrAlaCysAlaGlyGlyThrGlyCysAlaThrGlyCysCysAlaCys 1156
QY 237 CACCTCCAGGGCTCCTTGCCTTAGTGGCTGTAGCATCCCTACCAAGGAGGAGTGGT 178
Db 1157 CysAlaThrGlyCysCysAla---GlyGlyCysThrAlaAlaThrThrThrThrGly 1175
QY 177 GCGAATGACACAACTCAAGTTCGGAGGGGAACA---GGGAAGGAGGGATGGATGGGGGT 121
Db 1176 GlyThrGlyThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 1195
QY 120 GGTGTATCTTACTCTCTGTTTAAAGCAGAACACCTTGTTCGGGTGTTGGGAACA 70
Db 1196 GlyThrThrThrThrGly-----CysCysAlaThrGlyThr 1207
```

## RESULT 13

```
US-10-184-634-107
; Sequence 107, Application US/10184634
```

```
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 107
; LENGTH: 2478
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-184-634-107

Alignment Scores:
Pred. No.: 69 Length: 2478
Score: 73.50 Matches: 29
Percent Similarity: 35.05% Conservative: 5
Best Local Similarity: 29.90% Mismatches: 48
Query Match: 6.65% Indels: 15
DB: 9 Gaps: 4

US-09-914-191-1 (1-598) x US-10-184-634-107 (1-2478)
QY 357 TGCCTTGCTGGGATATCATCGGGATGCTTCTGGAAGGAACATCCAGATCTTAG 298
Db 1125 CysThrCysAlaAlaGlyThrAlaGlyCysThrGlyGly-----1137
QY 297 ATGAAGACACGTGACCCCTTTTGGCTTCAGTTCGTGGTGTCACTTAAACATCAAGCCCTCC 238
Db 1138 ---GlyAlaCysThrAlaCysAlaGlyGlyThrGlyCysAlaThrGlyCysAlaCys 1156
QY 237 CACCTCCAGGGCTCTTCGCTTAGTGGTGTAGCATCCCTACCCAGGACACTGGT 178
Db 1157 CysAlaThrGlyCysCysAla---GlyGlyCysThrAlaAlaThrThrThrThrGly 1175
QY 177 GCGAATGACACAACCTCAAGTTGGAGGGGAACA---GGGAAGGAAGGATGGATGGGGT 121
Db 1176 GlyThrGlyThrThrThrThrThrThrThrThrThrAlaGlyAlaGlyAlaCysThrGlyGly 1195
QY 120 GGTGATCTTACTCTGTTTAAAGCAGACACCTTGTTCGGGTGTTGGAACA 70
Db 1196 GlyThrThrThrThrThrGly-----CysCysAlaThrGlyThr 1207

RESULT 14
US-09-896-738-18
; Sequence 18, Application US/09896738
; Patent No. US20020165347A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Michael
; APPLICANT: Sullivan, John K.
; APPLICANT: Fang, Mei
; TITLE OF INVENTION: B7-Like Molecules and Uses Thereof
; FILE REFERENCE: 00-513-A
; CURRENT APPLICATION NUMBER: US/09/896,738
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,645
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 18

; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 107
; LENGTH: 2478
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-184-634-107

Alignment Scores:
Pred. No.: 69 Length: 2478
Score: 73.50 Matches: 29
Percent Similarity: 35.05% Conservative: 5
Best Local Similarity: 29.90% Mismatches: 48
Query Match: 6.65% Indels: 15
DB: 9 Gaps: 4

US-09-914-191-1 (1-598) x US-09-896-738-18 (1-584)
QY 74 CCAACACCCGCAACAAAGGTGTTCTGCTTAAACAGAGTAAGATACACACCCCATCCATC 133
Db 173 ProGlnProGlnIleIleYsTrpSer-----AspThrLysGlyGluAsnIle 187
QY 134 CCTTCCTCCCTGTTCCCTCCCAACTTGAGTTGTGTCAATCGCACCAAGTGTCTCTGGGTG 193
Db 188 ProAlaValGluAlaProVal-----ValAlaAspGly 198
QY 194 GTAGGGATGCTAC-----AGCCACCTAAGGCAAGGAGCCCTGGAGGTGGAG 241
Db 199 ValGlyLeu-TyrAlaValAlaAlaSerValIleMetArgGlySerSerGlyGlyGlyVa 218
QY 242 GGTTCGATGTTTAAAGCAGACACAGAACTGAAGCGCAAAAGGGTGCAGCTGCTTCTCATCTAG 301
Db 218 lSerCysIleIleIleArgAsnSerLeuLeuGlyLeuGlyLys-----ThrAlaSerIleSe 236
QY 302 AATCTCTGGATGTTCTCCAGAAAGCATCCCGATGATATCGCATGCAAGGGCACT 359
Db 236 rIleAlaAspProPheArgSerAlaGlnProTrpIleAlaAlaLeuAlaGlyThr 255

RESULT 15
US-09-910-174A-16
; Sequence 16, Application US/09910174A
; Patent No. US20020106730A1
; GENERAL INFORMATION:
; APPLICANT: Coyle, Anthony J.
; APPLICANT: Fraser, Christopher C.
; TITLE OF INVENTION: B7-H2 Molecules, No. US20020106730A1el Members of the B7
; FILE OF INVENTION: Family and Uses Thereof
; FILE REFERENCE: 35800/236924
; CURRENT APPLICATION NUMBER: US/09/910,174A
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 09/620,461
; PRIOR FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 584
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-910-174A-16

Alignment Scores:
Pred. No.: 51.8 Length: 584
Score: 73.00 Matches: 28
Percent Similarity: 48.00% Conservative: 20
Best Local Similarity: 28.00% Mismatches: 31
Query Match: 6.66% Indels: 21
DB: 10 Gaps: 4

US-09-914-191-1 (1-598) x US-09-910-174A-16 (1-584)
QY 74 CCAACACCCGCAACAAAGGTGTTCTGCTTAAACAGAGTAAGATACACACCCCATCCATC 133
Db 173 ProGlnProGlnIleIleYsTrpSer-----AspThrLysGlyGluAsnIle 187
QY 134 CCTTCCTCCCTGTTCCCTCCCAACTTGAGTTGTGTCAATCGCACCAAGTGTCTCTGGGTG 193
```

```
Db      |||:|:| |||:|:| |||:|:|
188 ProAlaValGluAlaProVal-----ValAlaAspGly 198
Qy      |||:|:| |||:|:| |||:|:|
194 GTAGGGATGCTAC-----AGCCACCTAAGCGAAGAGCCCTGGAGGTGGAG 241
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
199 ValGlyLeu-TyrAlaValAlaAlaSerValIleMetArgGlySerSerGlyGlyVa 218
Qy      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
242 GGCTTGCATGTTAAGCACACCACTGAAGCGCAAAGGGTCAGCTGTCTTCATCTAG 301
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
218 lSerCysIleIleArgAsnSerLeuLeuGlyLeuGluLys-----ThrAlaSerIleSe 236
Qy      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
302 AATCTCTGGATGTTCTTCAGAAAGCATCCCGATGATATCGACGTGCAAGGGCACT 359
Db      |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
236 rIleAlaAspProPhePheArgSerAlaGlnProTrrIleAlaAlaLeuAlaGlyThr 255
```

Search completed: July 9, 2003, 12:53:09  
Job time : 39 secs



QY 475 AGGAAGGAGCTGTTGATATGCGAAGGTCGATGAAGTATGACAAAGAAATTAAG 416  
 Db LysLysGlySerIleTrpIleArgGluProIleAspGlnGluTrpHisGlnGlnAspLys 225  
 QY 415 CTGCAACAGAAATGGAAGAAAGAAAGATGCGAGTACCCGGACCGAGCAAAAGCCAGTG 356  
 Db LysPheSerValLysGlnLysLysLysMetValLeuSer----- 238  
 QY 355 CCCTTGACATGCGATATCATCGGGGATGCTTTCTGGAAGGAACATCCA-----GAGATT 302  
 Db LysLysGlySerIleTrpIleArgGluProIleAspGlnGluTrpHisGlnGlnAspLys 225  
 QY 301 CTA 299  
 Db 259 Leu 259

## RESULT 2

S64315  
 hypothetical protein YGR024c - yeast (Saccharomyces cerevisiae)  
 N;Alternate names: hypothetical protein G4041  
 C;Species: Saccharomyces cerevisiae  
 C;Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 19-Apr-2002  
 C;Accession: S64315  
 R;Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.  
 submitted to the Protein Sequence Database, May 1996  
 A;Reference number: S64071  
 A;Accession: S64315  
 A;Molecule type: DNA  
 A;Residues: 1-237 <RIE>  
 A;Cross-references: EMBL:Z72809; NID:gl322995; PID:e243373; PID:gl322996; GSPDB:GN00007;  
 A;Experimental source: strain S288C  
 C;Genetics:

A;Gene: MIPS:YGR024c  
 A;Cross-references: SGD:S0003256  
 A;Map position: 7R  
 C;Superfamily: Saccharomyces cerevisiae hypothetical protein YGR024c

Alignment Scores:  
 Pred. No.: 0.000189 Length: 237  
 Score: 122.00 Matches: 23  
 Percent Similarity: 75.00% Conservative: 10  
 Best Local Similarity: 52.27% Mismatches: 11  
 Query Match: 11.04% Indels: 0  
 DB: 2 Gaps: 0

## US-09-914-191-1 (1-598) x S64315 (1-237)

QY 589 GGACTAACACAGTACAGGCCAAGGAGGATTAACAGGAACTCTTNCAGCAGACAAGAA 530  
 Db GlyLeuThrProGlnGluSerGluLysLeuCysGlyThrPheSerAsnGluLysGln 191  
 QY 529 GAGATTTTGTCTGAATTCACACTCAACTATAATAATGAGCTGCCGATGTATAGAAA 470  
 Db GluIleLeuPheSerGluCysGlyIleAsnTyrAsnAsnGluProGluMetPheLysLys 211  
 QY 469 GGGACTGTGTG 458  
 Db 212 GlySerLeuVal 215

## RESULT 3

G84731  
 hypothetical protein At2g32330 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 17-May-2002  
 C;Accession: G84731  
 R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Umayam, L.; Tallon, L.  
 Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A;Reference number: A84420; MUID:20083487; PMID:10617197  
 A;Accession: G84731

A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-297 <STO>  
 A;Cross-references: GB:AE002093; NID:g3831463; PIDN:AAC69945.1; GSPDB:GN00139  
 C;Genetics:  
 A;Gene: At2g32330  
 A;Map position: 2  
 C;Superfamily: Saccharomyces cerevisiae hypothetical protein YGR024c

Alignment Scores:  
 Pred. No.: 0.182 Length: 297  
 Score: 94.00 Matches: 20  
 Percent Similarity: 62.75% Conservative: 12  
 Best Local Similarity: 39.22% Mismatches: 19  
 Query Match: 8.51% Indels: 0  
 DB: 2 Gaps: 0

## US-09-914-191-1 (1-598) x G84731 (1-297)

QY 595 CAATCTGGACTAACACACAGGAGGATTAACAGGAACTCTTNCAGCAGAC 536  
 Db LysSerGlyLeuSerLysIleGlnAlaGlnAspTyrLeuLysGlyThrGlnThrArgGlu 246  
 QY 535 AAGAATGAGATTTGTTTCTGAATTCACACTCAACTATAATAATGAGCTGCCGATGTAT 476  
 Db LysAsnGluLeuLeuSerGlnGlnPheGlyIleGluTyrAsnSerLeuProValIlePhe 266  
 QY 475 AGGAAGGAGCTGTGTGATATGCGAAGGAGTG 443  
 Db 267 ArgMetGlySerSerValPheArgLeuLysVal 277

## RESULT 4

E84722  
 hypothetical protein At2g31580 [imported] - Arabidopsis thaliana  
 C;Species: Arabidopsis thaliana (mouse-ear cress)  
 C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C;Accession: E84722

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Umayam, L.; Tallon, L.  
 Nature 402, 761-768, 1999  
 A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A;Reference number: A84420; MUID:20083487; PMID:10617197  
 A;Accession: E84722  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-457 <STO>  
 A;Cross-references: GB:AE002093; NID:g4582470; PIDN:AAD24854.1; GSPDB:GN00139  
 C;Genetics:  
 A;Gene: At2g31580  
 A;Map position: 2

Alignment Scores:  
 Pred. No.: 0.506 Length: 457  
 Score: 90.00 Matches: 20  
 Percent Similarity: 57.14% Conservative: 12  
 Best Local Similarity: 35.71% Mismatches: 24  
 Query Match: 8.14% Indels: 0  
 DB: 2 Gaps: 0

## US-09-914-191-1 (1-598) x E84722 (1-457)

QY 595 CAATCTGGACTAACACACAGGAGGATTAACAGGAACTCTTNCAGCAGAC 536  
 Db LysSerGlyLeuSerLysIleGlnAlaGlnAspTyrLeuLysGlyThrGlnThrArgGlu 413  
 QY 535 AAGAATGAGATTTGTTTCTGAATTCACACTCAACTATAATAATGAGCTGCCGATGTAT 476  
 Db LysAsnGluLeuLeuSerGlnGlnPheGlyIleGluTyrAsnSerLeuProValIlePhe 433  
 QY 475 AGGAAGGAGCTGTGTGATATGCGAAGGAGTGATGACA 428  
 Db 434 ArgMetGlySerSerValPheArgLeuLysValSerIleAlaLeuThr 449



Db 69 ArgProGluLysLysGlnSerAspLysSerAsnTyrAlaArgAlaGluLeuPheArgGly 88  
 QY 370 TGGTCCGGTCA----- 381  
 Db 89 LysSerGlySerValSerPheAsnGlyLeuThrHisGlnLeuValGluGluSerLysLeu 108  
 QY 381 ----- 381  
 Db 109 ValSerAlaProPheGlnGluLysGlySerPheLeuTrpValLeuAlaProValVal 128  
 QY 382 -----CTGCCATCTTTTCTCCATCTTCCTCCATCTTCCTGGCAGCTTAATTT 423  
 Db 129 LeuLeuSerSerLeuLeuLeuProGlnPheLeuSerGlyIleLeuGluAlaThrPhe 148  
 QY 424 CTTTGTGCATCTTCATCCACCTTCGCCATATCAACA-----CAGTCCCTT 471  
 Db 149 LysAsnAspThrValAlaGlyArgSerGluIleValThrSerPheCysPheGluThrVal 168  
 QY 472 TCCTATACATCGGCAGCTCATTTATATAGT 501  
 Db 169 PheTyrAlaGlyLeuAlaIlePheLeuSer 178  
 RESULT 8  
 F84550  
 hypothetical protein At2g17310 [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: F84550  
 M.; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
 euss, D.; Niemman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: F84550  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-370 <SFO>  
 A:Cross-references: GB:AE002093; NID:q2262162; PIDN:AAB86508.1; GSPDB:GN00139  
 C:Genetics:  
 A:Gene: At2g17310  
 A:Map position: 2  
 C:Superfamily: Arabidopsis thaliana hypothetical protein F24M12.210  
 Alignment Scores:  
 Pred. No.: 3.95 Length: 370  
 Score: 81.50 Matches: 47  
 Percent Similarity: 38.12% Conservatives: 30  
 Best Local Similarity: 23.27% Mismatches: 68  
 Query Match: 7.38% Indels: 57  
 DB: 2 Gaps: 10  
 US-09-914-191-1 (1-598) x F84550 (1-370)  
 QY 520 TTTTCTGAATTCACATCACTATAATATGAGTGCCTCGC----- 482  
 Db 134 PheCysValPheGlyLeuGlyTyrAspAsnSerArgProGluLysValTyrLysIleLeu 153  
 QY 481 -----ATGTATAGGAAGGAGTGTGTTGATATGCAGAGGTG-----GATGAAGTG 434  
 Db 154 GlyHisLeuPheCysHisGlyLysValLeuArgAspGlnLysValIleTyrGluCys 173  
 QY 433 ATGACAAAGAAATTAAGCTGCCACAGAAATCGAAGGAAAAAGATGGCAGTGCAGCCGG 374  
 Db 174 AlSerAspSerLeuArgPheIleAspArgProGluAspAspTrpProIleThrGlu 193  
 QY 373 ACCGAGCAAGGACAGTGCCTTCGACTGCATATCATCGGGATGCTTCTCG----- 320  
 Db 194 ThrAlaLysArgSer-----AsnValSerLeuAsnGlyAsnLeuTyrTrpPheGly 210  
 QY 319 -----AAGGAACATCCAGAGATTCAGAT-----GAAGAC 290  
 Db 211 CysSerAsnTyrGluAsnAspGluTyrTyrIleArgIlePheAspPheSerThrGluAsp 230

QY 289 AGCTGACCCCTTTTGGCTTCAGTCTCGTGTGCTTAACCATGACGACCCCTCCACCTCCC 230  
 Db 231 PheLysProPheCys-----LeuLeu-ProCysGlnMetSerHisSerTh 245  
 QY 229 AGGGCTCCCTGGCTTAGGT-----GGCTGTAG 203  
 Db 245 rAspGluLeuValLeuAlaValTyrLysGlyAspArgPheSerLeuLeuLysGlnCysSe 265  
 QY 202 CATCCCT-----ACCACCAGGACACTGGTGGCAATGACACAACTCA 161  
 Db 265 rValThrArgGluIleGlyValTrpValThrLysGluArgIleSerAsnAsp----- 282  
 QY 160 AGTTGGGAGGGACAGGAGGAGGATGGGTGGGTGGTGTATCTTACTCTCTTTA 101  
 Db 283 -----AsnGlyAsnGlyGlyGluGlu-TrpLeuLysLeuMetThrLeuSerL 300  
 QY 100 AGCAGAAC-----ACCTTGTTCGGTGTGGAAATCATGTTCTCTTGGCAGAGTCTT 47  
 Db 300 ysProAsnLeuProLysLeuPheGlyThrValSerTyrPheIleTyrGlyLysThrLeuT 320  
 QY 46 TT 45  
 Db 320 yr 320  
 RESULT 9  
 T51779  
 non-phototropic hypocotyl 3-like protein - Arabidopsis thaliana  
 N;Alternate names: protein F28D10.10  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: T51779  
 R.; Delzeny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Rudd, S.;  
 submitted to the Protein Sequence Database, August 2000  
 A:Reference number: Z25454  
 A:Accession: T51779  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-661 <DEL>  
 A:Cross-references: EMBL:AL391254  
 A:Experimental source: cultivar Columbia; BAC clone F28D10  
 C:Genetics:  
 A:Map position: 3  
 A:Introns: 28/3; 83/3; 419/3  
 A:Note: F28D10\_10  
 C:Superfamily: Arabidopsis hypothetical protein F19F18.80  
 Alignment Scores:  
 Pred. No.: 4.21 Length: 661  
 Score: 81.50 Matches: 45  
 Percent Similarity: 41.34% Conservatives: 29  
 Best Local Similarity: 25.14% Mismatches: 62  
 Query Match: 7.38% Indels: 43  
 DB: 2 Gaps: 9  
 US-09-914-191-1 (1-598) x T51779 (1-661)  
 QY 505 ATCACTATAATAATGAGTCCGATGTATAGAAAGGACTGTG----- 461  
 Db 398 LeuAsnTyrSerAspSerGluThrLeuTyrAsnValAspCysValGluArgIleValArg 417  
 QY 460 TTGATATGCACAAAGTGGATCAAGTGCACAAA-----GAAATTAAGCTG--- 413  
 Db 418 HisPheTrpArgLeuValAspSerTyrMetAlaGluValAlaSerAspValAsnLeuLys 437  
 QY 412 CCAACAGAAATGAAGGAAAAAGATGGCAGTGACCCGACGAGCAAGGACGAGTCCC 353  
 Db 438 ProAspLysMetArgSerLeuAlaAlaLeuProGluSerSerArg-----Pro 454  
 QY 352 TTGCACTGGCATATCATCGG-----GATGCTTCTGGAGCAACATCCAGATTTCTA 299  
 Db 455 LeuTyrAspGlyLeuTyrArgAlaPheAspIleTyrPheLysGluHisProTrpLeuSer 474



Qy 298 GATGAACACACCTGACCCCTTTTGC----- 275  
 Db 475 AspArgAspGlySerGluGlnLeuCyAsnIleMetAspTyrGlnArgLeuSerIleAspAla 494  
 Qy 274 -----GCTTCA-----GTTCTGGTGGCTTAACC 251  
 Db 495 CysAlaHisAlaSerHisAsnAspArgLeuProLeuArgValValLeuGlnValLeuPhe 514  
 Qy 250 ATGCAAGCCCTCCACCTCCAGGGCTCCTTAGTGCTGGCTGTAGCATCCCTACCAC 191  
 Db 515 PheGluGlnMet-HisLeuArgThrAlaLeuAla--GlyGlyLeuAsnValAlaAsnTh 533  
 Qy 190 CCAGGACACTGGTCCGAATGACACAACTCAAGTTGGAGGGGACACGGGAAGAGGAT 131  
 Db 533 rGIuThr-----AlaHisAlaValThrIleProGlyGlyArgThrGlyGlnGluIleVa 551  
 Qy 130 GGATGGGGTGGTGTATCTTACTCTGTTTAAGCAGACACCTTGTTCGGGTG 78  
 Db 551 iGlnArgAspGly-TrpValThrValValArgGlnAsnGlnValLeuLysVal 568  
 RESULT 10  
 JE0211  
 hypothetical 44k protein - Bombyx mori cytoplasmic polyhedrosis virus strain I  
 C;Species: Bombyx mori cytoplasmic polyhedrosis virus, BmCPV  
 C;Date: 21-Aug-1998 #sequence\_revision 21-Aug-1998 #text\_change 20-Jun-2000  
 C;Accession: JE0211  
 R;Hagiwara, K.; Tomita, M.; Kobayashi, J.; Miyajima, S.; Yoshimura, T.  
 Biochem. Biophys. Res. Commun. 247, 549-553, 1998  
 A;Title: Nucleotide sequence of Bombyx mori cytoplasmic polyhedrosis virus segment 8.  
 A;Reference number: JE0211; MUID:98321164; PMID:9657006  
 A;Accession: JE0211  
 A;Molecule type: mRNA  
 A;Residues: 1-390 <HAG>  
 A;Cross-references: GB:AB016436; NID:g3395150; PIDN:BAA32042.1; PID:g3395151  
 Alignment Scores:  
 Pred. No.: 4.49 Length: 390  
 Score: 81.00 Matches: 27  
 Percent Similarity: 49.02% Conservative: 23  
 Best Local Similarity: 26.47% Mismatches: 44  
 Query Match: 7.33% Indels: 8  
 DB: 2 Gaps: 4

US-09-914-191-1 (1-598) x JE0211 (1-390)  
 Qy 577 GTCAAGCCCAAGGAGGATACAGGAACTCTTNCAGCAGACAGAATGAGATTTTCTTT 518  
 Db 136 IIEGluSerGluGlyGluValAsnAspThrProGlnAspAspGluAlaGluLysLysAsp 155  
 Qy 517 TCTGAATTCAACTCAACTAATAAT-----GAGCTGCCGATGTATAGGAAAGGG 467  
 Db 156 ValGluIleArgMetAsnTrpSerGluGluIleValGluLeuProArgGlnGluGlu 175  
 Qy 466 ACTGTGTGTATATGCGAGAGGTGGATGAGTGTATGATGACAAAGAAATAGTCCCAACA 407  
 Db 176 AsnValLeuValLeuSerLysProSerMetIleSerGluGluGluLeu---MetProThr 194  
 Qy 406 GAAATGAAGGAAAAAGATGGCAGTGCACCGGACAGGACAAAGCCAGTGCCTTCAC 347  
 Db 195 AspMetGlu-----ValAlaThrProLysValLeuGluProProThrProLeuPro 211  
 Qy 346 TGCATATCATCGGGATGCTTTCTGGAAGGAACATCCAGAGATTTCTAGATGAGACAGC 287  
 Db 212 SerProIleIle---ValAlaValSerSerGluSerProGlnValLysGluIleGluArg 230  
 Qy 286 TGACCC 281  
 Db 231 ProPro 232  
 RESULT 11  
 AF2172  
 membrane-bound lytic transglycosylase A alr2933 [imported] - Nostoc sp. (strain PCC 7120)  
 C;Species: Nostoc sp.

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C;Accession: AF2172  
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuguch  
 N.; Kaneko, T.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,  
 DNA Res. 8, 205-213, 2001  
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An  
 A;Reference number: AB1807; MUID:21595285; PMID:11759840  
 A;Accession: AF2172  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-410 <KUR>  
 A;Cross-references: GB:BA000019; PIDN:BA074632.1; PID:g17132027; GSPDB:GN00179  
 A;Experimental source: strain PCC 7120  
 C;Genetics:  
 A;Gene: alr2933  
 Alignment Scores:  
 Pred. No.: 5.1 Length: 410  
 Score: 80.50 Matches: 44  
 Percent Similarity: 35.33% Conservative: 15  
 Best Local Similarity: 26.35% Mismatches: 44  
 Query Match: 7.29% Indels: 64  
 DB: 2 Gaps: 7

US-09-914-191-1 (1-598) x AF2172 (1-410)  
 Qy 478 TATAGGAAGGGGACTGTGTGATATGCGACAAGGTGGATGAAGTGATGACAAAGAAATT 419  
 Db 251 TyrAlaGlyGlyThrAspTyrProThrThrSerIleGlyArgGluLeuAlaLysAspGly 270  
 Qy 418 AAGCTGCCA---ACAGAAATGGAAGAAAAAGATGGCAGTGCACCCGGACAGGACAAAG 362  
 Db 271 LysLeuProLeuSerGlyMetThrMetProLysMetIle----- 283  
 Qy 361 CCACTGCCCTTGACATGCGATATCATCGGGGATGCTTCTGGGAAGAACATCCAGAGATT 302  
 Db 284 -----SerPhePheArgGlnGlnProGluGlu 292  
 Qy 301 CTA-----GATGAAGACAGACTGA 284  
 Db 293 LeuAsnAsnTyrLeuProArgTrpGluArgPheValPhePheGlnGluThrGlySerArg 312  
 Qy 283 CCCTTTTGCCTTCAGTCTCTGCTGCTTAACCATGCAA----- 245  
 Db 313 ProAlaThrGlySerIleAsnValProValThrProGluArgSerIleAlaThrAspLys 332  
 Qy 244 GCCTCTCCACCTCCAGGGCTCCTTCCTTAGTGGCTGTAGCATCCCT----- 196  
 Db 333 SerLeuMetPro-ProGlyAlaLeuAlaLeuValHisThrSerIleProPheProValAl 352  
 Qy 195 -----ACCACCCAGGACACTGGTGC 176  
 Db 352 aGlyArgMetGluTyrArgThrValSerArgTyrValLeuAspGlnAspThrGly--- 371  
 Qy 175 GAATGACACAACTCAAGTTGGGAGG-----GGAAACAGGAAGGAAGGAT 131  
 Db 372 -AsnAlaIleLysGlyProGlyArgValAspTyrPheMetGlyThrGlyLysGlnAlaGl 391  
 Qy 130 GGATGGGGTGGTGTATCT 112  
 Db 391 yAspArgAlaGlyValThr 397  
 RESULT 12  
 T18319  
 hypothetical protein L7610.10 - Leishmania major  
 C;Species: Leishmania major  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 11-Jan-2002  
 C;Accession: T18319  
 R;Oliver, K.; Murphy, L.; Quail, M.; Lawson, D.; Harris, D.; Rajandream, M.; Ivens, A.;  
 submitted to the EMBL Data Library, May 1999  
 A;Reference number: Z18876  
 A;Accession: T18319

A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-320 <OLI>  
 A;Cross-references: EMBL:AL034356; NID:e1371878; PIDN:CAA22245.1  
 C;Genetics:  
 A;Note: L7610.10

## Alignment Scores:

Pred. No.: 7.17 Length: 320  
 Score: 79.00 Matches: 44  
 Percent Similarity: 31.98% Conservative: 11  
 Best Local Similarity: 25.58% Mismatches: 52  
 Query Match: 7.21% Indels: 65  
 DB: 2 Gaps: 7

US-09-914-191-1 (1-598) x T18319 (1-320)

Qy 78 CACGCAACAAAGGTGTTCTGCTTAAACAGAGTAAGATACAC---CACCCCAT----- 128  
 Db 69 HisArgLeuCySerLeuLeuSerLeuLeuLeuHisThrHisProHisIleThr 88  
 Qy 128 ----- 128  
 Db 89 TyrValHisThrSerThrArgAsnHisLeuSerLeuCySerValPheLeuSerGlyLysAsp 108  
 Qy 129 -----CCATCCCTTCTTCCCTGTTCCCTCC 155  
 Db 109 AlaThrSerValArgAlaProSerProProSerProProSerLeuHisProser 128  
 Qy 156 CAACCTGAGTTGTGTCATTCGCACCACTGCTCGGTGGTAGGATGCTACGCCACCTA 215  
 Db 129 ArgLeu-----CysTyr----- 132  
 Qy 216 AGGCAAGAGCCCTGGAGGTGGAGGGCTTGC-----ATGGTTAAGCACACAGAACT 269  
 Db 133 -----GlyArgCysArgSerAlaAlaSerHisLeuAlaGlu 145  
 Qy 270 GAAGCGCAAAGGTGAGTCTTCATCTAGATCTCTGGATCTTCTCCAGAAAGCA 329  
 Db 146 AlaGlyArgGlyGlnIleAlaArgGluArgArgLysGlyThrGlyArgGAsp 165  
 Qy 330 TCCCGCATGATATCGCAGTGCAGGGCACTGGCTTGTCTCTGGTCC-----GGGTC 380  
 Db 166 GlyProSerProProSerCysAlaCysPheSerPheValLeu-SerValAlaGluThrTh 185  
 Qy 381 ACTGCCATCTTTTCTTCATTCGTTGGCAGCTTAATTTCTTTGTCATCACTTCA 440  
 Db 185 rLeuProProMetArgSerPro-----SerAlaProProLeuLeuSerProPr 203  
 Qy 441 TCCACCTTCTGCCATATCAACACAGTCCCTTCC 474  
 Db 203 oProProSerSerThrProProSerGlnLeuAla 214

## RESULT 13

T31614

hypothetical protein Y50E8A.j - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C;Accession: T31614

R;Steward, C.

submitted to the EMBL Data Library, September 1999

A;Reference number: 221047

A;Accession: T31614

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-225 &lt;WIL&gt;

A;Cross-references: EMBL:AL117200; NID:e1549770; PIDN:CAB55053.1; CESP:Y50E8A.j

A;Experimental source: clone Y50E8A

C;Genetics:

A;Gene: CESP:Y50E8A.j

## Alignment Scores:

Pred. No.: 7.8 Length: 225

Score: 78.50 Matches: 33  
 Percent Similarity: 37.86% Conservative: 20  
 Best Local Similarity: 23.57% Mismatches: 37  
 Query Match: 7.16% Indels: 50  
 DB: 2 Gaps: 7

US-09-914-191-1 (1-598) x T31614 (1-225)

Qy 220 AAGAGCCCTGGAGGTGGGA-----GGGCTTCATGG 252  
 Db 109 LysHisProTrpAspProGlyThrGlnGlnArgTyrGlnLeuAlaProArgThrGlnTrp 128  
 Qy 253 TTAAGCACACCAGAACTGAAGCGCAAGGGTCAGCTGCTTCATCTAGAACTCTCTGGAT 312  
 Db 129 MetProThrProLeuLeuAspGlnAlaAlaIle-LeuSerPheLeuSerIleLeuTyrIle 148  
 Qy 313 GTTCCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGGGCACTGGCTTTGTCCTGG 372  
 Db 148 ephe-----SerIleProHis-----G1 154  
 Qy 373 TCGGGTCACTGCCATCTTTTCTTCCATTTCTCTTGCAGCTT-----AA 420  
 Db 154 yProSerThrCysAsnValPheSerIleTrpProGlnTrpAsnPheGlyAsnSerAsnAs 174  
 Qy 421 TTTCCTT-----TGTCACTCACTTCCACCTTCTGCCATATCAACA 462  
 Db 174 nTyrTrpMetProIleGlyGlnGlySerHisPheSerTyrLeuLeuProGlnSe 194  
 Qy 463 CAGTCCCTTTCTATACATCGGAGCTCATATTATATAGTTGATGTGAATTGAGAAACA 522  
 Db 194 xAla-----AlaAsnLeuGlnPheArgLeuPr 203  
 Qy 523 AAATCTCATCTTGTCTGC-----TGNAAGAGTTCCTCTGAATCTCCC 565  
 Db 203 oGluPheSerMetIleCysTyrGlnArgTyrGlnIleLysSerThrMetAspCysPro 222

## RESULT 14

H84071

lipopolysaccharide biosynthesis BH3376 [imported] - Bacillus halodurans (strain C-125)

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C;Accession: H84071

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Fuji, F.; Hir;

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A;Reference number: A83650; WUID:20512592; PMID:11058132

A;Accession: H84071

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-397 &lt;STO&gt;

A;Cross-references: GB:AP001518; GB:BA000004; NID:G10175792; PIDN:BAB07095.1; GSPDB:GNO

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH3376

## Alignment Scores:

Pred. No.: 8.29 Length: 397  
 Score: 78.50 Matches: 35  
 Percent Similarity: 40.98% Conservative: 15  
 Best Local Similarity: 28.69% Mismatches: 35  
 Query Match: 7.16% Indels: 37  
 DB: 2 Gaps: 8

US-09-914-191-1 (1-598) x H84071 (1-397)

Qy 19 TATAAATAAGTACTCGGATTAAAAAAGCACTT-----CTG 57

Db 149 TrpLysTyrSerThrGluPheGluSerArgAlaLeuSerGlnHisIleIleThr 168

Qy 58 CCAAGGAACCATGTTCCACACCCGCAACAGGTGTTCTGCTTAACACAGATAATAC 117

Db 169 ProSerAsnTrpLeuAsnHisTyrLeuLysGluLeuPheCysLeuLysGlyValArgIle 188

QY 118 ACCACCCCATCCATCCCTTCCTTCCTGTTCC----- 150  
 Db 189 ThrThr---ValHisAsnGlyLeuSerValSerSerPheLeuSerLysLeuAsnGlnAla 207  
 QY 151 -----CCTCCCAAC-----TTGAGTTCTGTCATTCGCACAGT 183  
 Db 208 PheHisProSerSerGlnSerLysLysValLeuSerCysIleAlaArgLeuThr 227  
 QY 184 GTCCCTGGTGGTAGG-----GATGCTACAGCCACCTAAGCAAGGAG- 225  
 Db 228 ProLeuLysGlyHisIleTyrLeuLeuAspAlaLeuAlaGluLeuLysLysAsnThrLeu 247  
 QY 226 CCTCGGAGGTGGAGGGCTTGCATGG---TTAAGCACACAGAACTGAAGCGCAAAAGG 282  
 Db 248 AspTrpGlu-----CysTrpPheIleGlyAsnGlyGluIleLysLysLeu 263  
 QY 283 GTCAGC 288  
 Db 264 ValAsn 265

RESULT 15  
 T18274  
 1-phosphatidylinositol 3-kinase (EC 2.7.1.137) 3 - slime mold (Dictyostelium discoideum)  
 C;Species: Dictyostelium discoideum  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T18274  
 R;Zhou, K.; Takegawa, K.; Emr, S.D.; Firtel, R.A.  
 Mol. Cell. Biol. 15, 5645-5656, 1995  
 A;Title: A phosphatidylinositol (PI) kinase gene family in Dictyostelium discoideum: Bio  
 A;Reference number: 206411  
 A;Accession: T18274  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-1585 <ZHO>  
 A;Cross-references: EMBL:U23478; NID:G733523; PID:G733524; PIDN:AAA85723.1  
 C;Genetics:  
 A;Gene: PIK3  
 C;Keywords: phosphotransferase

Alignment Scores:  
 Pred. No.: 9.62 Length: 1585  
 Score: 78.50 Matches: 51  
 Percent Similarity: 36.99% Conservative: 30  
 Best Local Similarity: 23.29% Mismatches: 62  
 Query Match: 7.10% Indels: 76  
 DB: 2 Gaps: 10

US-09-914-191-1 (1-598) x T18274 (1-1585)

QY 571 GCCCAAGGGAGATTACAGGAACTCTTNCAGCACAGAAATGAGATTTGTTTCT--- 515  
 Db 498 AlaSerGlyLysMetGlnGlyHisLeuAsnLeuGluLysAspGluLysPheThrLeu 517  
 QY 514 -----GAA 512  
 Db 518 ArgTrpCysAsnLysAspValPheAspGlnAspThrProLeuGlyHisLeuLeuGln 537  
 QY 511 TTCAACATCACTATATATAGCTGCCGATGTATAGGAAGGGACTGTGTGATATGG 452  
 Db 538 TyrAsnLeuAsnTyrAsnAsn-----ProThrGlnLysProThrAsnIle----- 552  
 QY 451 CAGAAGGTGATGAGTGAACAAAGAAATTAAGCTGCCAACAGAAATGGAGGAAAA 392  
 Db 553 ---LysLeuGluLeuValLeuGluAspGluLeuCysLysGluArgLeuValAspLeuGln 571  
 QY 391 AAGATGCAGTACCCGACGAGCAACAGCCAGTGCCTTGCACCTCCGATATCATCGGG 332  
 Db 572 SerLeuGluIleAsnAsnGlyArg----- 579  
 QY 331 GATGCTTTCTGGAGGAGACAT---CCAGATTTCTAGATGAAGACAGCTGACCTTTTGC 275  
 Db 580 ProSerIleTrpLysSerHisIleAspValLeuSerPheAsn----- 594

QY 274 GCTTCAGTTCTGTGTGCTTAACCATGTCAAGGCCTCCACCTCCAGGGCTCCTTCGCTT 215  
 Db 595 ---ArgLysLeuArgGluLeuAlaMetLeuAlaLysProGlnSerAsnValProAlaAla 613  
 QY 214 AGGTGGCTGTAGCATCCCTACACCCAGGACACTGGTGGCAATGACACAACCTCAAGTTGG 155  
 Db 614 ArgLeuThrProTyrPro--ProProLysThrIleProGluPhePhe----- 628  
 QY 154 GAGGGGAACAGGGAAGGAGGTGGATGGGGTGGTGTATCTTACTCTGTTTAAAG---C 98  
 Db 629 -----ValIleArgValHisLeuPheLysAsnG 638  
 QY 97 AGAACACCTTGTGTTGGGTGGTGGAAAC-----ATGGTTCCTTTG 59  
 Db 638 ln-ThrLysSerLeuArgCysAlaAsnAsnHisThrAlaPheSerLeuMetThrIleLeu 657  
 QY 58 GCAGAGTCTT-----TTTTTTTAAATCGCAGTACTATTTT 23  
 Db 658 SerGluLysLeuLysAsnThrThrProPheAspProThrGlnTyrArgPhe 674

Search completed: July 9, 2003, 12:51:24  
 Job time : 30 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 9, 2003, 12:45:33 ; Search time 13 Seconds  
(without alignments)  
3815.822 Million cell updates/sec

Title: US-09-914-191-1  
Perfect score: 1096  
Sequence: 1 ttggaatagttcttctgctta.....ggtgttagccagattgttg 598

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p\_model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_epool/US09914191/runat\_09072003\_112348\_12866/app\_query.fasta\_1.775  
-DB=SwissProt\_40 -QFMT=fastan -SUFFIX=rs -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09914191 @CGN 1.1 26 @runat\_09072003\_112348\_12866 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESOURCY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	122	11.0	237	1 YG1G YEAST	P53215 saccharomyc
2	85.5	7.8	495	1 ENP2 RAT	O35795 rattus norv
3	84.5	7.7	495	1 ENP2 MOUSE	O55026 mus musculu
C 4	80	7.2	304	1 PUR7 PICPA	Q3C114 pichia past
5	80	7.3	673	1 FXO3 HUMAN	O43524 homo sapien
6	79.5	7.3	2144	1 CLR2 RAT	Q9GY72 rattus norv
C 7	78.5	7.1	1585	1 P3K3 DICDI	P54675 dictyosteli
C 8	78	7.1	700	1 DTNB MOUSE	O70585 mus musculu
C 9	78	7.1	743	1 DTNA HUMAN	O9Y418 homo sapien
C 10	78	7.1	862	1 SLA2 BACAN	P94217 bacillus an
C 11	78	7.1	2907	1 FEN2 MOUSE	Q61555 mus musculu
C 12	77	7.0	627	1 DTNB HUMAN	O60941 homo sapien
C 13	76	6.9	304	1 GAL PSEFL	P11886 pseudomonas
C 14	76	6.9	490	1 MVOG MOUSE	O70624 mus musculu
C 15	76	6.9	1570	1 P3K1 DICDI	P54673 dictyosteli
C 16	75.5	6.9	512	1 FEN2 YEAST	P25621 saccharomyc
17	75.5	6.9	615	1 CPO DROME	Q01617 drosophila
18	75	6.8	415	1 YLN2 CAEEL	Q18964 caenorhabdi

C 19	74	6.7	469	1 FXGA HUMAN	P55316 homo sapien
20	73.5	6.7	544	1 YN62 TREPA	O83932 treponema p
21	73.5	6.7	591	1 MNT MOUSE	O08789 mus musculu
C 22	73	6.6	502	1 MYOC RAT	O08789 mus musculu
C 23	73	6.6	3010	1 POLG HCVJA	O92318 rhizobium m
C 24	72.5	6.6	551	1 AGLA RHIME	P26662 h genome po
C 25	72.5	6.6	1013	1 PTPX MACNE	O22378 rhizobium m
C 26	72.5	6.6	3010	1 POLG HCVTW	P29846 h genome neme
C 27	72	6.6	141	1 YPRO OWEFU	P21260 owenia fusi
28	72	6.6	684	1 SKIL HUMAN	P12757 homo sapien
C 29	71.5	6.5	395	1 HOC1 YEAST	P47124 saccharomyc
C 30	71.5	6.5	611	1 ILVD AGRTS	O8ue43 agrobacteri
C 31	71	6.4	247	1 APMI MOUSE	O60994 mus musculu
C 32	71	6.4	331	1 ODPB PORPU	P51266 porphyra pu
C 33	71	6.5	614	1 NRDI HUMAN	P20393 homo sapien
C 34	71	6.5	1473	1 OVOS CHICK	P20740 gallus gall
35	71	6.5	2911	1 FBN2 HUMAN	P35556 homo sapien
36	71	6.5	2923	1 CLR2 HUMAN	O9hcu4 homo sapien
C 37	70.5	6.4	306	1 PUR7 YEAST	P27616 saccharomyc
C 38	70.5	6.4	478	1 CYCA GLUOX	O47945 gluconobact
C 39	70.5	6.4	1124	1 POL_FIVT2	P31822 feline immu
C 40	70.5	6.4	3010	1 POLG HCVBK	P26663 h genome po
C 41	70.5	6.4	3010	1 POLG HCVJT	O00269 h genome po
C 42	70	6.3	291	1 PUR7 CANNA	P27602 candida mal
43	70	6.4	386	1 SH1B SPAEH	P56496 spalax leuc
44	70	6.4	478	1 OPN4 HUMAN	O9uhm6 homo sapien
C 45	70	6.3	776	1 TFR1 CHICK	O90997 gallus gall

ALIGNMENTS

RESULT 1  
YG1G YEAST STANDARD; PRT; 237 AA.  
AC PS3215;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 27.8 kDa protein in VMA7-RPS25A intergenic region.  
GN YGR024C.  
OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C;  
RX MEDLINE=97435481; PubMed=9290212;  
RX Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;  
RT "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae  
chromosome VII.";  
RL Yeast 13:1077-1090(1997)  
CC -!- SIMILARITY: TO S.POMBE SPCC63.07.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Z72809; CAA97007.1;  
DR SGD: S0003256; YGR024C.  
DR Hypothetical protein.  
SQ SEQUENCE 237 AA; 27757 MW; CCB018E24B9EAE26 CRC64;

Alignment Scores:  
Pred. No.: 0.000128 Length: 237  
Score: 122.00 Matches: 23  
Percent Similarity: 75.00% Conservative: 10  
Best Local Similarity: 52.27% Mismatches: 11  
Query Match: 11.04% Indels: 0



Db 363 ValGlyThrLeuLysGlnLeuGlu-----GluAlaThrGluIleThrCysAsnGlnThr 380

QY 567 TGG 569

Db 381 Trp 381

# RESULT 3

## ENP2\_MOUSE

ID ENP2\_MOUSE STANDARD; PRT; 495 AA.  
AC O55026; O35928; Q9DCR9;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ectonucleoside triphosphate diphosphohydrolase 2 (EC 3.6.1.3)  
GN ENTPD2 (Ecto-ATPase) (CD39 antigen-like 1).  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (SHORT FORM).  
RC TISSUE=Embryo;  
RX MEDLINE=97419269; PubMed=9271669;  
RA Chadwick B.P., Frischauf A.-M.;  
RT "Cloning and mapping of a human and mouse gene with homology to ecto-ATPase genes.";  
RL Mamm. Genome 8:668-672(1997).  
[2]  
RP SEQUENCE FROM N.A. (LONG FORM).  
RC TISSUE=Hepatoma;  
RX MEDLINE=98288263; PubMed=9624117;  
RA Gao L., Dong L., Whitlock J.P. Jr.;  
RT "A novel response to dioxin. Induction of ecto-ATPase gene expression.";  
RL J. Biol. Chem. 273:15358-15365(1998).  
[3]  
RP SEQUENCE FROM N.A. (LONG FORM).  
RC STRAIN=C57BL/6J; TISSUE=Kidney;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Azawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio I.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baren G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.;  
RT Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
CC -!- FUNCTION: IN THE NERVOUS SYSTEM, COULD HYDROLYZE ATP AND OTHER  
CC NUCLEOTIDES TO REGULATE PURINERGIC NEUROTRANSMISSION. HYDROLYZES  
CC ADP ONLY TO A MARGINAL EXTENT (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate;  
CC -!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- INDUCTION: BY DIOXIN.  
CC -!- PTM: HAS PROBABLY DISULFIDE BONDS.  
CC -!- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----

DR EMBL; U91511; AAB81014.1; -;  
DR EMBL; AF042811; AAC24347.1; -;  
DR EMBL; AK002553; BAB22182.1; -;  
DR MGD; MGI-1096863; Entpd2  
DR InterPro; IPR000407; GDAL\_CD39\_NTPase.  
DR Pfam; PF01150; GDAL\_CD39; 1.  
DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.  
KW Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;  
KW Alternative splicing.  
FT DOMAIN 1 4 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 5 25 POTENTIAL.  
FT DOMAIN 26 462 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 463 483 POTENTIAL.  
FT DOMAIN 484 495 CYTOPLASMIC (POTENTIAL).  
FT CARBOHYD 64 64 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 129 129 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 294 294 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 319 319 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 378 378 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 443 443 N-LINKED (GLCNAC...) (POTENTIAL).  
FT VARSPLIC 130 132 LTS -> MAG (IN SHORT ISOFORM).  
FT VARSPLIC 133 495 MISSING (IN SHORT ISOFORM).  
FT CONFLICT 400 400 T -> A (IN REF. 3).  
FT CONFLICT 414 414 S -> R (IN REF. 3).  
FT CONFLICT 437 437 T -> A (IN REF. 3).  
SQ SEQUENCE 495 AA; 54310 MW; CC0811C5D79CA4C8 CRC64;

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
87	84.50	495	42
Percent Similarity:	37.80%	Conservative:	20
Best Local Similarity:	25.61%	Mismatches:	69
Query Match:	7.71%	Indels:	33
DB:	1	Gaps:	9

US-09-914-191-1 (1-598) x ENP2\_MOUSE (1-495)

QY	87	CAAGGTGTTCTGCTTAAACAGAGTAGATACACACCCATCCATCCCTTCCTCCCTG	146
Db	250	GlnArgLeuLeuAlaSerAlaLeuGlnIleHisArgPheHisProCysTrpProLysGly	269
QY	147	TTCCCTCCCAACTTGAGTTGTTCATTCGCACCAAGTGTCTCTGGTGGTAGGATCTAC	206
Db	270	TyrSerThrGlnValLeuLeuArgGluValTyrGlnSerPro-----CysThr	285
QY	207	AGCCACTAAGCAAGGAGCCCTGGAGGTGGAGGCTTGATGGTTAAGCACACACAGA	266
Db	286	MetGlyGlnArgProGlnThrPheAsnSerSerAlaThrValSerLeuSerGlyThrSer	305
QY	267	ACTGAAGCCCAAAAGGTGAGTGTCTTCATCT-----AGAATCTCTGATGTTCTCTC	320
Db	306	AsnAlaAlaLeuCysArgAspLeuValSerGlyLeuPheAsnIleSerSerCysProPhe	325
QY	321	CAGAAAGCATCCCGATGATATCGCAGTGCAGAGGCATGTGGCTTTGTCTGTCGGGTC	380
Db	326	-----SerGlnCysSerPheAsnGly-----	332
QY	381	ACTGCCATCTTTTCTTCATTTCTGTGTGCAGCTTAATTTCTTTTGTCACTTCA	440
Db	333	-----ValPheGlnProProValAla---GlyAsnPheIleAlaPhe-----	345
QY	441	TCCACCTCTGCGCATATCAACAGTCCCTTTCTTATACATCGGAGCTCATTTATTATAG	500
Db	346	SerAlaPheTyrTyr-----ThrValAspPheLeu-----LysThrValMetGly	360
QY	501	TTGATGTTGAATTCAGAAACAAAATCTCATTTCTGTCTGCTGNAAGAGTTCCTCGTAAT	560

Db 361 LeuProValGlyThrLeuLysGlnLeuGlu-----AspAlaThrGluThrThrCysAsn 378  
 Qy 561 CTCCTTGGGCT 572  
 Db 379 GlnThrTrpAla 382

## RESULT 4

PUR7\_PICPA STANDARD; PRT; 304 AA.  
 AC Q9C1G4;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Phosphoribosylamidimidazole-succinocarboxamide synthase (EC 6.3.2.6)  
 DE (SAICAR synthetase).  
 GN ADE1.  
 OS Pichia pastoris (Yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Pichia.  
 OX NCBI\_TaxID=4922;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21125227; PubMed=1123254;  
 RA Lin Cereghino G.P., Lin Cereghino J., Sunga A.J., Johnson M.A.,  
 RA Lim M., Gleeson M.A.G., Clegg J.M.;  
 RT "New selectable marker/auxotrophic host strain combinations for  
 RT molecular genetic manipulation of Pichia pastoris.";  
 RL Gene 263:159-169(2001).  
 CC -1- CATALYTIC ACTIVITY: ATP + 5-amino-1-(5-phospho-D-  
 CC ribosyl)imidazole-4-carboxylate + L-aspartate = ADP + phosphate +  
 CC (S)-2-[5-amino-1-(5-phospho-D-ribosyl)imidazole-4-  
 CC carboxamido]succinate.  
 CC -1- PATHWAY: De novo purine biosynthesis; seventh step.  
 CC -1- SIMILARITY: BELONGS TO THE SAICAR SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
 CC -----  
 CC EMBL; AF321096; AA06766.1; -  
 DR HSP; P27616; IA48.  
 DR InterPro; IPR001636; SAICAR\_synt.  
 DR Pfam; PF01259; SAICAR\_synt; 1.  
 DR ProDom; PD003043; SAICAR\_synt; 1.  
 DR TIGRfam; TIGR00081; purG; 1.  
 DR PROSITE; PS01057; SAICAR\_SYNTHETASE\_1; 1.  
 DR PROSITE; PS01058; SAICAR\_SYNTHETASE\_2; 1.  
 KW Ligase; Purine biosynthesis.  
 SQ SEQUENCE 304 AA; 34101 MW; 584073DA712262C9 CRC64;

## Alignment Scores:

Pred. No.: 3.03 Length: 304  
 Score: 80.00 Matches: 29  
 Percent Similarity: 40.94% Conservative: 23  
 Best Local Similarity: 22.83% Mismatches: 37  
 Query Match: 7.24% Indels: 38  
 DB: 1 Gaps: 6

US-09-914-191-1 (1-598) x PUR7\_PICPA (1-304)

Qy 589 GGACTACACAGTACAGCCAGGAGGATTACAGGAACTCTTCAGCAGACAGAAT 530  
 Db 10 GlyileLeuProLeulleAlaLysGlyLysAlaValArgAspIleTyrlAlaValAspGluAsn 29  
 Qy 529 GAGATTGTTGTT-----TCTGAATTCACATCACTATAATAGAG 488  
 Db 30 AsnLeuLeuPheValAlaThrAspArgIleSerAlaTyrlAspValIleMetThrAsnGly 49

Qy 487 CTGCCGATGTATAGGAAGGACTGTGTTGATATGCAGAGGTG----- 443  
 Db 50 IlePro-----AspLysGlyLysIleLeuThrGlnLeuSerValPheThrPheAspPhe 67  
 Qy 442 -----GATGAAGTGTGACA--- 428  
 Db 68 LeuAlaProTyrlIleLysAsnHisLeuValAlaSerAsnAspLysGluValPheAlaLeu 87  
 Qy 427 -----AAAGAAATTAACTGCCACAGAAATGGAAGGAAAAAGATGGCA 383  
 Db 88 LeuProSerLysLeuSerGluGluTyrlLysSerGlnLeuGluGlyArgSerLeuIle 107  
 Qy 382 GTGACCCGACAGCAGCAAGCCAGTGCCTTGCCTGCACTGATATCATCGG----- 332  
 Db 108 ValLysLysHisArgLeu-----IleProLeuGluAlaIleValArgGlyTyrlIleThr 125  
 Qy 331 GATGCTTTCTGGAAGCAACAT 311  
 Db 126 GlySerAlaTrpLysGluTyr 132  
 RESULT 5  
 FX03\_HUMAN STANDARD; PRT; 673 AA.  
 ID FX03\_HUMAN  
 AC O43524; O15171;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Forkhead box protein O3A (Forkhead in rhabdomyosarcoma-like 1) (AF6q21  
 DE protein).  
 GN FOXO3A OR FKHL1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98140118; PubMed=9479491;  
 RA Anderson M.J., Viars C.S., Czekay S., Cavenee W.K., Arden K.C.;  
 RT "Cloning and characterization of three human forkhead genes that  
 RT comprise an FKHR-like gene subfamily.";  
 RL Genomics 47:187-199(1998).  
 CC [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RA Strausberg R.;  
 RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 CC [3]  
 RP SEQUENCE OF 1-383 FROM N.A.  
 RX MEDLINE=98008138; PubMed=9345057;  
 RA Hillion J., Le Coniat M., Jonveaux P., Berger R., Bernard O.A.;  
 RT "AF6q21, a novel partner of the MLL gene in t(6;11)(q21;q23), defines  
 RT a forkhead transcriptional factor subfamily.";  
 RL Blood 90:3714-3719(1997).  
 CC -1- FUNCTION: PROBABLE TRANSCRIPTION FACTOR THAT MAY TRIGGER  
 CC APOPTOSIS BY INDUCING THE EXPRESSION OF GENES THAT ARE CRITICAL  
 CC FOR CELL DEATH.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS.  
 CC -1- PTM: PHOSPHORYLATED BY AKT1.  
 CC -1- DISEASE: INVOLVED IN A T(6;11)(Q21;Q23) CHROMOSOMAL TRANSLOCATION  
 CC IN SECONDARY ACUTE LEUKEMIA THAT INVOLVES FOXO3A AND MLL.  
 CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.  
 CC -1- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;  
 CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/AF6q21ID125.html".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).  
 CC -----



```

DR EMBL; AF032886; AAC39592.1; --
DR EMBL; BC020227; AAH20227.1; --
DR EMBL; BC021224; AAH21224.1; --
DR EMBL; AJ001589; CAA04860.1; --
DR EMBL; AJ001590; CAA04861.1; --
DR HSSP; O63245; 2HFH.
DR TRANSFAC; T02938; --
DR Genew; HGNC:3821; FOXO3A.
DR MIN; 602681; --
DR InterPro; IPR001766; TF_Fork_head.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PF00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR SMART; SM00339; FH; 1.
DR PROSITE; PS00657; FORK_HEAD_1; FALSE_NEG.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS0039; FORK_HEAD_3; 1.
KW Transcription regulation; DNA-binding; Nuclear protein; Apoptosis;
KW Chromosomal translocation; Proto-oncogene; Phosphorylation.
FT DNA BIND 157 251 FORK-HEAD.
FT CONFLICT 156 163 AWGNLSYA -> WGNPVYS (IN REF. 3).
FT CONFLICT 238 246 PDGKSGKA -> LMGERKT (IN REF. 3).
FT CONFLICT 253 253 S -> T (IN REF. 3).
FT CONFLICT 271 271 MISSING (IN REF. 3).
FT CONFLICT 292 330 PGSPTRSSDLDLAWTFDRKTSNASTVSGRLSPIMAS
-> AWQPHVNAAVMSWMRGRTSVHAPILTPAQSVAAACRPWS
QV (IN REF. 3).
FT CONFLICT 345 361 PMLYSSASLSPSVSKP -> AHALQHSQPVTFSKQA
(IN REF. 3).
FT CONFLICT 367 367 P -> R (IN REF. 3).
FT CONFLICT 371 371 D -> E (IN REF. 3).
FT CONFLICT 382 383 LT -> AD (IN REF. 3).
SQ SEQUENCE 673 AA; 71276 MW; ESB4E830665A9982 CRC64;

Alignment Scores:
Pred. No.: 3.43 Length: 673
Score: 80.00 Matches: 27
Percent Similarity: 40.40 Conservative: 13
Best Local Similarity: 27.27 Mismatches: 26
Query Match: 7.30 Indels: 33
DB: 1 Gaps: 2

US-09-914-191-1 (1-598) x FOXO3_HUMAN (1-673)
QY 87 CAAGGTGTTCTGCTTAACAGAGTACATACACACCCCATCCCTTCCTTCCTG 146
Db 525 GlnGlySerLeuValAsnGlnAsnLeuHisGlnHisGlnThr----- 540
QY 147 TTCCCTCCCAACTTGAGTTGTGTCTCATTCGCACACAGTGTCTGGGTAGGATGCTAC 206
Db 540 ----- 540
QY 207 AGCCACCTAAGGCAAGAGCCCTGGGAGGTGGGAGGCTTCATGTTAAGCACACAGA 266
Db 541 -----GlnGlyAlaLeuGlyGlySerArgAlaLeu-----SerAsn 552
QY 267 ACTGAAGGCCAAAGAGTGTCTCTCATCTAGAAATCTGTGATGTTCTCTCCAGAAA 326
Db 553 SerValSerAsnMetGlyLeuSerGluSerSerLeuGlySerAlaLeuHisGlnGln 572
QY 327 GCATCCCGGATGATATCCAGTGCAGAGGCATGGCTTGTCTCTGGTCCGGGTCA 381
Db 573 GlnSerProVal-SerGlnSerMetGlnThrLeuSerAspSerLeuSerGlySer 590

RESULT 6
CLR2_RAT
ID CLR2_RAT STANDARD; PRT; 2144 AA.
AC Q9QY2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cadherin EGF LAG seven-pass G-type receptor 2 (Multiple epidermal
growth factor-like domains 3) (Fragment).

```

```

GN CELSR2 OR MEGF3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=98360089; PubMed=9693030;
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT Identification of high-molecular-weight proteins with multiple
RT EGF-like motifs by motif-trap screening.;
RL Genomics 51:27-34(1998).
CC -1- FUNCTION: Receptor that may have an important role in cell/cell
CC signaling during nervous system formation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in the brain. High expression in
CC cerebellum and olfactory bulb. Weaker expression in cerebral
CC cortex, hippocampus and brain stem.
CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- SIMILARITY: CONTAINS 4 CADHERIN DOMAINS.
CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB011529; BAA8687.1; --
DR HSSP; P00740; LEDM.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR002126; Cadherin.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR000832; GPCR secretin.
DR InterPro; IPR001879; horxn receptor.
DR InterPro; IPR002049; Laminin EGF.
DR InterPro; IPR001791; Laminin G.
DR InterPro; IPR000203; PKC_cys_rich.
DR Pfam; PF00002; 7tm 2; 1.
DR Pfam; PF00028; cadherin; 3.
DR Pfam; PF01825; GPS; 6.
DR Pfam; PF02793; HRM; 1.
DR Pfam; PF00054; laminin G; 1.
DR PRINTS; PR00205; CADHERIN.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGFLAMININ.
DR PRINTS; PR00249; GPCRSECRETIN.
DR SMART; SM00112; CA; 3.
DR SMART; SM00179; EGF_CA; 1.
DR SMART; SM00001; EGF_like; 6.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; HormR; 1.
DR SMART; SM00282; LamG; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00232; CADHERIN_1; 3.
DR PROSITE; PS00268; CADHERIN_2; 4.
DR PROSITE; PS00221; GPS; 1.
DR PROSITE; PS00025; LAM_G_DOMAIN; 2.
DR PROSITE; PS00022; EGF_1; 6.
DR PROSITE; PS01186; EGF_2; 4.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR PROSITE; PS00227; G_PROTEIN_RECEP_F2_3; 1.

```



RA Zhou K., Takegawa K., Emr S.D., Firtel R.A.;  
 RT "A phosphatidylinositol (PI) kinase gene family in Dictyostelium  
 RT discoidum: biological roles of putative mammalian p110 and yeast  
 RL Mol. Cell. Biol. 15:5645-5656(1995)  
 CC -!- CATALYTIC ACTIVITY: ADP + 1-phosphatidyl-ID-myo-inositol = ADP +  
 CC 1-phosphatidyl-ID-myo-inositol 3-phosphate.  
 CC -!- SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or  
 CC send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; U23478; AAA85723.1; -  
 DR Dictyob; DD01101; pikC.  
 DR InterPro; IPR002420; PI3K\_C2.  
 DR InterPro; IPR000341; PI3K\_ras\_bind.  
 DR InterPro; IPR001263; PI3Ka.  
 DR InterPro; IPR000403; PI3\_P14\_kinase.  
 DR Pfam; PF00454; PI3\_P14\_kinase; 1.  
 DR Pfam; PF00613; PI3Ka; 1.  
 DR Pfam; PF00792; PI3K\_C2; 1.  
 DR Pfam; PF00794; PI3K\_rbd; 1.  
 DR SMART; SM00142; PI3K\_C2; 1.  
 DR SMART; SM00144; PI3K\_rbd; 1.  
 DR SMART; SM00145; PI3Ka; 1.  
 DR SMART; SM00146; PI3K; 1.  
 DR PROSITE; PS00915; PI3\_4\_KINASE\_1; 1.  
 DR PROSITE; PS00916; PI3\_4\_KINASE\_2; 1.  
 DR PROSITE; PS0290; PI3\_4\_KINASE\_3; 1.  
 DR TRANSFERASE; Kinase; Multigene family; Repeat.  
 FT NON\_TER 1 1  
 FT DOMAIN 58 84 POLY-ASN.  
 FT DOMAIN 98 110 POLY-ASN.  
 FT DOMAIN 200 226 POLY-ASN.  
 FT DOMAIN 239 254 POLY-ASN.  
 FT DOMAIN 345 378 POLY-ASN.  
 FT DOMAIN 383 390 POLY-ASN.  
 FT DOMAIN 484 488 POLY-SER.  
 FT DOMAIN 720 737 POLY-GLN.  
 FT DOMAIN 1569 1575 POLY-ASN.  
 FT DOMAIN 1221 1484 PI3K/PI4K.  
 FT DOMAIN 1510 1539 5 X 5 AA APPROXIMATE REPEATS.  
 FT REPEAT 1510 1514 1.  
 FT REPEAT 1515 1519 2.  
 FT REPEAT 1520 1524 3.  
 FT REPEAT 1530 1534 4.  
 FT REPEAT 1535 1539 5.  
 FT DOMAIN 1547 1560 7 X 2 AA TANDEM REPEATS OF K-E.  
 SQ SEQUENCE 1585 AA; 180421 MW; 4689B620D2484961 CRC64;  
 Alignment Scores:  
 Pred. No.: 5.61 Length: 1585  
 Score: 78.50 Matches: 51  
 Percent Similarity: 36.99% Conservative: 30  
 Best Local Similarity: 23.29% Mismatches: 62  
 Query Match: 7.10% Indels: 76  
 DB: 1 Gaps: 10  
 US-09-914-191-1 (1-598) x PI3K\_DICD1 (1-1585)  
 QY 571 GCCAAGGAGATACAGGAACTCTTNCAGCAGACAGAATGAGATTTTGTCTTCT--- 515  
 Db 498 AlaserglyLysMetGlnGlyHisLeuAsnLeuGluLysAspGluLysPhePheThrLeu 517  
 QY 514 -----GAA 512  
 Db 518 ArgTrpCysAsnLysAspValValPheAspGlnAspThrProLeuGlyHisLeuLeuGln 537

QY 511 TTCAACATCACTATAATATGAGCTGCCGATGATATAGAAAGGAGGACTGTGTGATATGG 452  
 Db 538 TyrAsnLeuAsnTyrAsnAsn-----ProThrGlnLysProThrAsnLeu----- 552  
 QY 451 CAGAAGGTGGATGAAGTATGACAAAGAAATTAAGTCCCAACACAAATGGAAGGAAAA 392  
 Db 553 ---LysLeuGluLeuValLeuGluAspGluLeuCysLysGluArgLeuValAspLeuGln 571  
 QY 391 AAGATGGCAGTACCCGGACAGGACCAAGCCAGTGCCTTGTGCACTGCGATCATATCGGG 332  
 Db 572 SerLeuGluLeuAsnAsnGlyArg----- 579  
 QY 331 GATGCTTTCTGGAAGAACAT---CCAGAGATTTCTAGTGAAGACAGCTGACCTTTTTC 275  
 Db 580 ProSerIleTrpLysSerHisLeuAspValLeuSerPheAsn----- 594  
 QY 274 GCTTCAGTTCTGTGTGCTTAACCATGCAAGCCCTCCACCTCCAGGGCTCCTTGCCTT 215  
 Db 595 ---ArgLysLeuArgGluLeuAlaMetLeuAlaLysProGlnSerAsnValProAlaAla 613  
 QY 214 AGGTGGCTGTAGCATCCCTACACCCAGGACACTGGTGGCAATGACACAACTCAAGTTGG 155  
 Db 614 ArgLeuThrProTyrPro--ProProLysThrIleProGluPhePhe----- 628  
 QY 154 GAGGGGAACAGGGAAGGATGGATGGGGTGTGTATCTTCTTCTTCTTCTTCTTCTTCT 98  
 Db 629 -----ValIleArgValHisLeuPheLysAsnG 638  
 QY 97 AGAACACCTTGTGGGTGGTGGAC-----ATGGTTCCTTTC 59  
 Db 638 In-ThrLysSerLeuArgCysAlaAsnAsnHisThrAlaPheSerLeuMetThrLeuLeu 657  
 QY 58 GCAGAAAGTGTCTT-----TTTTTAAATCGCAGTACTATTTT 23  
 Db 658 SerGluLysLeuLysAsnThrThrProPheAspProThrGlnTyrArgPhe 674  
 RESULT 8  
 DTNB\_MOUSE STANDARD; PRT; 700 AA.  
 AC 070585; 070563; Q9CTZ1;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Dystrobrevin beta (Beta-dystrobrevin) (DTN-B)  
 GN DTNB.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=98200066; PubMed=9540997;  
 RA Puca A.A., Piluso V.N.G., Belsito A., Sampaolo S., Quaderi N.,  
 RA Rossi E., Di Iorio G., Ballabio A., Franco B.;  
 RT "Identification and characterization of a novel member of the  
 RT dystrobrevin gene family.";  
 RL FEBS Lett. 425:7-13(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=98081858; PubMed=9419360;  
 RA Blake D.J., Nawrotzki R., Loh N.Y., Gorecki D.C., Davies K.E.;  
 RT "Beta-dystrobrevin, a member of the dystrophin-related protein  
 RT family.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 95:241-246(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2).  
 RX MEDLINE=99018217; PubMed=9799833;  
 RA Loh N.Y., Ambrose H.J., Guay-Woodford L.M., Dasgupta S.,  
 RA Nawrotzki R.A., Blake D.J., Davies K.E.;  
 RT "Genomic organization and refined mapping of the mouse beta-  
 RT dystrobrevin gene.";  
 RL Mamm. Genome 9:857-862(1998).  
 RN [4]



RT related phosphoprotein found at the Torpedo electric organ post-synaptic membrane.";  
 RL Hum. Mol. Genet. 5:489-496(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 3 AND 7).  
 RX MEDLINE=20197320; PubMed=10735273;  
 RA Sacculiet-Puccio H.M., Feener C.A., Schaid D.J., Thibodeau S.N.,  
 RA Michels V.V., Kunkel L.M.;  
 RT "The genomic organization of human dystrobrevin.";  
 RL Neurogenetics 1:37-42(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 7).  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=98369056; PubMed=9701558;  
 RA Nawrozi R., Loh N.Y., Ruegg M.A., Davies K.E., Blake D.J.;  
 RT "Characterisation of alpha-dystrobrevin in muscle.";  
 RL J. Cell Sci. 111:2595-2605(1998).  
 CC -!- FUNCTION: MAY BE INVOLVED IN THE FORMATION AND STABILITY OF SYNAPSES AS WELL AS BEING INVOLVED IN THE CLUSTERING OF NICOTINIC ACETYLCHOLINE RECEPTORS.  
 CC -!- SUBUNIT: INTERACTS WITH DYSTROPHIN, UTPROPHIN AND SYNTROPHIN. ISOFORMS 7 AND 8 DO NOT INTERACT WITH DYSTROPHIN.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -!- ALTERNATIVE PRODUCTS: AT LEAST 8 ISOFORMS; 1/DTN-1 (SHOWN HERE), 2/DYSTROBREVIN-ALPHA, 3/DTN-2, 4/DYSTROBREVIN-BETA, 5/DYSTROBREVIN-GAMMA, 6/DYSTROBREVIN-EPSILON, 7/DTN-3/ALPHA-DYSTROBREVIN-3/DYSTROBREVIN-DELTA AND 8/DYSTROBREVIN-ZETA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN, SKELETAL AND CARDIAC MUSCLES, AND EXPRESSED AT LOWER LEVELS IN LUNG, LIVER AND PANCREAS. ISOFORM 2 IS NOT EXPRESSED IN CARDIAC MUSCLE. ISOFORMS 7 AND 8 ARE ONLY EXPRESSED IN MUSCLE.  
 CC -!- DOMAIN: THE COILED-COIL DOMAIN MEDIATE THE INTERACTION WITH DYSTROPHIN AND UTPROPHIN (BY SIMILARITY).  
 CC -!- PTM: PHOSPHORYLATION OF DTN-1 ON TYROSINE KINASE SUBSTRATE DOMAIN PRESENT IN THE C-TERMINUS (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE DYSTROPHIN FAMILY. DYSTROBREVIN SUBFAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U46744; AAC50429.1; -;  
 DR EMBL; U46745; AAC50430.1; -;  
 DR EMBL; U26744; AAC50426.1; -;  
 DR EMBL; U46746; AAC50431.1; -;  
 DR EMBL; U26742; AAC50424.1; -;  
 DR EMBL; U26743; AAC50425.1; -;  
 DR EMBL; U46744; AAC50429.1; -;  
 DR EMBL; U46745; AAC50430.1; -;  
 DR EMBL; U46746; AAC50431.1; -;  
 DR EMBL; U46747; AAC50432.1; -;  
 DR EMBL; U46748; AAC50433.1; -;  
 DR EMBL; U46749; AAC50434.1; -;  
 DR EMBL; U46750; AAC50435.1; -;  
 DR EMBL; U46751; AAC50436.1; -;  
 DR EMBL; U46752; AAC50437.1; -;  
 DR EMBL; U46753; AAC50438.1; -;  
 DR EMBL; U46754; AAC50439.1; -;  
 DR EMBL; U46755; AAC50440.1; -;  
 DR EMBL; U46756; AAC50441.1; -;  
 DR EMBL; U46757; AAC50442.1; -;  
 DR EMBL; U46758; AAC50443.1; -;  
 DR EMBL; U46759; AAC50444.1; -;  
 DR EMBL; U46760; AAC50445.1; -;  
 DR EMBL; U46761; AAC50446.1; -;  
 DR EMBL; U46762; AAC50447.1; -;  
 DR EMBL; U46763; AAC50448.1; -;  
 DR EMBL; U46764; AAC50449.1; -;  
 DR EMBL; U46765; AAC50450.1; -;  
 DR EMBL; U46766; AAC50451.1; -;  
 DR EMBL; U46767; AAC50452.1; -;  
 DR EMBL; U46768; AAC50453.1; -;  
 DR EMBL; U46769; AAC50454.1; -;  
 DR EMBL; U46770; AAC50455.1; -;  
 DR EMBL; U46771; AAC50456.1; -;  
 DR EMBL; U46772; AAC50457.1; -;  
 DR EMBL; U46773; AAC50458.1; -;  
 DR EMBL; U46774; AAC50459.1; -;  
 DR EMBL; U46775; AAC50460.1; -;  
 DR EMBL; U46776; AAC50461.1; -;  
 DR EMBL; U46777; AAC50462.1; -;  
 DR EMBL; U46778; AAC50463.1; -;  
 DR EMBL; U46779; AAC50464.1; -;  
 DR EMBL; U46780; AAC50465.1; -;  
 DR EMBL; U46781; AAC50466.1; -;  
 DR EMBL; U46782; AAC50467.1; -;  
 DR EMBL; U46783; AAC50468.1; -;  
 DR EMBL; U46784; AAC50469.1; -;  
 DR EMBL; U46785; AAC50470.1; -;  
 DR EMBL; U46786; AAC50471.1; -;  
 DR EMBL; U46787; AAC50472.1; -;  
 DR EMBL; U46788; AAC50473.1; -;  
 DR EMBL; U46789; AAC50474.1; -;  
 DR EMBL; U46790; AAC50475.1; -;  
 DR EMBL; U46791; AAC50476.1; -;  
 DR EMBL; U46792; AAC50477.1; -;  
 DR EMBL; U46793; AAC50478.1; -;  
 DR EMBL; U46794; AAC50479.1; -;  
 DR EMBL; U46795; AAC50480.1; -;  
 DR EMBL; U46796; AAC50481.1; -;  
 DR EMBL; U46797; AAC50482.1; -;  
 DR EMBL; U46798; AAC50483.1; -;  
 DR EMBL; U46799; AAC50484.1; -;  
 DR EMBL; U46800; AAC50485.1; -;  
 DR EMBL; U46801; AAC50486.1; -;  
 DR EMBL; U46802; AAC50487.1; -;  
 DR EMBL; U46803; AAC50488.1; -;  
 DR EMBL; U46804; AAC50489.1; -;  
 DR EMBL; U46805; AAC50490.1; -;  
 DR EMBL; U46806; AAC50491.1; -;  
 DR EMBL; U46807; AAC50492.1; -;  
 DR EMBL; U46808; AAC50493.1; -;  
 DR EMBL; U46809; AAC50494.1; -;  
 DR EMBL; U46810; AAC50495.1; -;  
 DR EMBL; U46811; AAC50496.1; -;  
 DR EMBL; U46812; AAC50497.1; -;  
 DR EMBL; U46813; AAC50498.1; -;  
 DR EMBL; U46814; AAC50499.1; -;  
 DR EMBL; U46815; AAC50500.1; -;  
 DR EMBL; U46816; AAC50501.1; -;  
 DR EMBL; U46817; AAC50502.1; -;  
 DR EMBL; U46818; AAC50503.1; -;  
 DR EMBL; U46819; AAC50504.1; -;  
 DR EMBL; U46820; AAC50505.1; -;  
 DR EMBL; U46821; AAC50506.1; -;  
 DR EMBL; U46822; AAC50507.1; -;  
 DR EMBL; U46823; AAC50508.1; -;  
 DR EMBL; U46824; AAC50509.1; -;  
 DR EMBL; U46825; AAC50510.1; -;  
 DR EMBL; U46826; AAC50511.1; -;  
 DR EMBL; U46827; AAC50512.1; -;  
 DR EMBL; U46828; AAC50513.1; -;  
 DR EMBL; U46829; AAC50514.1; -;  
 DR EMBL; U46830; AAC50515.1; -;  
 DR EMBL; U46831; AAC50516.1; -;  
 DR EMBL; U46832; AAC50517.1; -;  
 DR EMBL; U46833; AAC50518.1; -;  
 DR EMBL; U46834; AAC50519.1; -;  
 DR EMBL; U46835; AAC50520.1; -;  
 DR EMBL; U46836; AAC50521.1; -;  
 DR EMBL; U46837; AAC50522.1; -;  
 DR EMBL; U46838; AAC50523.1; -;  
 DR EMBL; U46839; AAC50524.1; -;  
 DR EMBL; U46840; AAC50525.1; -;  
 DR EMBL; U46841; AAC50526.1; -;  
 DR EMBL; U46842; AAC50527.1; -;  
 DR EMBL; U46843; AAC50528.1; -;  
 DR EMBL; U46844; AAC50529.1; -;  
 DR EMBL; U46845; AAC50530.1; -;  
 DR EMBL; U46846; AAC50531.1; -;  
 DR EMBL; U46847; AAC50532.1; -;  
 DR EMBL; U46848; AAC50533.1; -;  
 DR EMBL; U46849; AAC50534.1; -;  
 DR EMBL; U46850; AAC50535.1; -;  
 DR EMBL; U46851; AAC50536.1; -;  
 DR EMBL; U46852; AAC50537.1; -;  
 DR EMBL; U46853; AAC50538.1; -;  
 DR EMBL; U46854; AAC50539.1; -;  
 DR EMBL; U46855; AAC50540.1; -;  
 DR EMBL; U46856; AAC50541.1; -;  
 DR EMBL; U46857; AAC50542.1; -;  
 DR EMBL; U46858; AAC50543.1; -;  
 DR EMBL; U46859; AAC50544.1; -;  
 DR EMBL; U46860; AAC50545.1; -;  
 DR EMBL; U46861; AAC50546.1; -;  
 DR EMBL; U46862; AAC50547.1; -;  
 DR EMBL; U46863; AAC50548.1; -;  
 DR EMBL; U46864; AAC50549.1; -;  
 DR EMBL; U46865; AAC50550.1; -;  
 DR EMBL; U46866; AAC50551.1; -;  
 DR EMBL; U46867; AAC50552.1; -;  
 DR EMBL; U46868; AAC50553.1; -;  
 DR EMBL; U46869; AAC50554.1; -;  
 DR EMBL; U46870; AAC50555.1; -;  
 DR EMBL; U46871; AAC50556.1; -;  
 DR EMBL; U46872; AAC50557.1; -;  
 DR EMBL; U46873; AAC50558.1; -;  
 DR EMBL; U46874; AAC50559.1; -;  
 DR EMBL; U46875; AAC50560.1; -;  
 DR EMBL; U46876; AAC50561.1; -;  
 DR EMBL; U46877; AAC50562.1; -;  
 DR EMBL; U46878; AAC50563.1; -;  
 DR EMBL; U46879; AAC50564.1; -;  
 DR EMBL; U46880; AAC50565.1; -;  
 DR EMBL; U46881; AAC50566.1; -;  
 DR EMBL; U46882; AAC50567.1; -;  
 DR EMBL; U46883; AAC50568.1; -;  
 DR EMBL; U46884; AAC50569.1; -;  
 DR EMBL; U46885; AAC50570.1; -;  
 DR EMBL; U46886; AAC50571.1; -;  
 DR EMBL; U46887; AAC50572.1; -;  
 DR EMBL; U46888; AAC50573.1; -;  
 DR EMBL; U46889; AAC50574.1; -;  
 DR EMBL; U46890; AAC50575.1; -;  
 DR EMBL; U46891; AAC50576.1; -;  
 DR EMBL; U46892; AAC50577.1; -;  
 DR EMBL; U46893; AAC50578.1; -;  
 DR EMBL; U46894; AAC50579.1; -;  
 DR EMBL; U46895; AAC50580.1; -;  
 DR EMBL; U46896; AAC50581.1; -;  
 DR EMBL; U46897; AAC50582.1; -;  
 DR EMBL; U46898; AAC50583.1; -;  
 DR EMBL; U46899; AAC50584.1; -;  
 DR EMBL; U46900; AAC50585.1; -;  
 DR EMBL; U46901; AAC50586.1; -;  
 DR EMBL; U46902; AAC50587.1; -;  
 DR EMBL; U46903; AAC50588.1; -;  
 DR EMBL; U46904; AAC50589.1; -;  
 DR EMBL; U46905; AAC50590.1; -;  
 DR EMBL; U46906; AAC50591.1; -;  
 DR EMBL; U46907; AAC50592.1; -;  
 DR EMBL; U46908; AAC50593.1; -;  
 DR EMBL; U46909; AAC50594.1; -;  
 DR EMBL; U46910; AAC50595.1; -;  
 DR EMBL; U46911; AAC50596.1; -;  
 DR EMBL; U46912; AAC50597.1; -;  
 DR EMBL; U46913; AAC50598.1; -;  
 DR EMBL; U46914; AAC50599.1; -;  
 DR EMBL; U46915; AAC50600.1; -;  
 DR EMBL; U46916; AAC50601.1; -;  
 DR EMBL; U46917; AAC50602.1; -;  
 DR EMBL; U46918; AAC50603.1; -;  
 DR EMBL; U46919; AAC50604.1; -;  
 DR EMBL; U46920; AAC50605.1; -;  
 DR EMBL; U46921; AAC50606.1; -;  
 DR EMBL; U46922; AAC50607.1; -;  
 DR EMBL; U46923; AAC50608.1; -;  
 DR EMBL; U46924; AAC50609.1; -;  
 DR EMBL; U46925; AAC50610.1; -;  
 DR EMBL; U46926; AAC50611.1; -;  
 DR EMBL; U46927; AAC50612.1; -;  
 DR EMBL; U46928; AAC50613.1; -;  
 DR EMBL; U46929; AAC50614.1; -;  
 DR EMBL; U46930; AAC50615.1; -;  
 DR EMBL; U46931; AAC50616.1; -;  
 DR EMBL; U46932; AAC50617.1; -;  
 DR EMBL; U46933; AAC50618.1; -;  
 DR EMBL; U46934; AAC50619.1; -;  
 DR EMBL; U46935; AAC50620.1; -;  
 DR EMBL; U46936; AAC50621.1; -;  
 DR EMBL; U46937; AAC50622.1; -;  
 DR EMBL; U46938; AAC50623.1; -;  
 DR EMBL; U46939; AAC50624.1; -;  
 DR EMBL; U46940; AAC50625.1; -;  
 DR EMBL; U46941; AAC50626.1; -;  
 DR EMBL; U46942; AAC50627.1; -;  
 DR EMBL; U46943; AAC50628.1; -;  
 DR EMBL; U46944; AAC50629.1; -;  
 DR EMBL; U46945; AAC50630.1; -;  
 DR EMBL; U46946; AAC50631.1; -;  
 DR EMBL; U46947; AAC50632.1; -;  
 DR EMBL; U46948; AAC50633.1; -;  
 DR EMBL; U46949; AAC50634.1; -;  
 DR EMBL; U46950; AAC50635.1; -;  
 DR EMBL; U46951; AAC50636.1; -;  
 DR EMBL; U46952; AAC50637.1; -;  
 DR EMBL; U46953; AAC50638.1; -;  
 DR EMBL; U46954; AAC50639.1; -;  
 DR EMBL; U46955; AAC50640.1; -;  
 DR EMBL; U46956; AAC50641.1; -;  
 DR EMBL; U46957; AAC50642.1; -;  
 DR EMBL; U46958; AAC50643.1; -;  
 DR EMBL; U46959; AAC50644.1; -;  
 DR EMBL; U46960; AAC50645.1; -;  
 DR EMBL; U46961; AAC50646.1; -;  
 DR EMBL; U46962; AAC50647.1; -;  
 DR EMBL; U46963; AAC50648.1; -;  
 DR EMBL; U46964; AAC50649.1; -;  
 DR EMBL; U46965; AAC50650.1; -;  
 DR EMBL; U46966; AAC50651.1; -;  
 DR EMBL; U46967; AAC50652.1; -;  
 DR EMBL; U46968; AAC50653.1; -;  
 DR EMBL; U46969; AAC50654.1; -;  
 DR EMBL; U46970; AAC50655.1; -;  
 DR EMBL; U46971; AAC50656.1; -;  
 DR EMBL; U46972; AAC50657.1; -;  
 DR EMBL; U46973; AAC50658.1; -;  
 DR EMBL; U46974; AAC50659.1; -;  
 DR EMBL; U46975; AAC50660.1; -;  
 DR EMBL; U46976; AAC50661.1; -;  
 DR EMBL; U46977; AAC50662.1; -;  
 DR EMBL; U46978; AAC50663.1; -;  
 DR EMBL; U46979; AAC50664.1; -;  
 DR EMBL; U46980; AAC50665.1; -;  
 DR EMBL; U46981; AAC50666.1; -;  
 DR EMBL; U46982; AAC50667.1; -;  
 DR EMBL; U46983; AAC50668.1; -;  
 DR EMBL; U46984; AAC50669.1; -;  
 DR EMBL; U46985; AAC50670.1; -;  
 DR EMBL; U46986; AAC50671.1; -;  
 DR EMBL; U46987; AAC50672.1; -;  
 DR EMBL; U46988; AAC50673.1; -;  
 DR EMBL; U46989; AAC50674.1; -;  
 DR EMBL; U46990; AAC50675.1; -;  
 DR EMBL; U46991; AAC50676.1; -;  
 DR EMBL; U46992; AAC50677.1; -;  
 DR EMBL; U46993; AAC50678.1; -;  
 DR EMBL; U46994; AAC50679.1; -;  
 DR EMBL; U46995; AAC50680.1; -;  
 DR EMBL; U46996; AAC50681.1; -;  
 DR EMBL; U46997; AAC50682.1; -;  
 DR EMBL; U46998; AAC50683.1; -;  
 DR EMBL; U46999; AAC50684.1; -;  
 DR EMBL; U47000; AAC50685.1; -;  
 DR EMBL; U47001; AAC50686.1; -;  
 DR EMBL; U47002; AAC50687.1; -;  
 DR EMBL; U47003; AAC50688.1; -;  
 DR EMBL; U47004; AAC50689.1; -;  
 DR EMBL; U47005; AAC50690.1; -;  
 DR EMBL; U47006; AAC50691.1; -;  
 DR EMBL; U47007; AAC50692.1; -;  
 DR EMBL; U47008; AAC50693.1; -;  
 DR EMBL; U47009; AAC50694.1; -;  
 DR EMBL; U47010; AAC50695.1; -;  
 DR EMBL; U47011; AAC50696.1; -;  
 DR EMBL; U47012; AAC50697.1; -;  
 DR EMBL; U47013; AAC50698.1; -;  
 DR EMBL; U47014; AAC50699.1; -;  
 DR EMBL; U47015; AAC50700.1; -;  
 DR EMBL; U47016; AAC50701.1; -;  
 DR EMBL; U47017; AAC50702.1; -;  
 DR EMBL; U47018; AAC50703.1; -;  
 DR EMBL; U47019; AAC50704.1; -;  
 DR EMBL; U47020; AAC50705.1; -;  
 DR EMBL; U47021; AAC50706.1; -;  
 DR EMBL; U47022; AAC50707.1; -;  
 DR EMBL; U47023; AAC50708.1; -;  
 DR EMBL; U47024; AAC50709.1; -;  
 DR EMBL; U47025; AAC50710.1; -;  
 DR EMBL; U47026; AAC50711.1; -;  
 DR EMBL; U47027; AAC50712.1; -;  
 DR EMBL; U47028; AAC50713.1; -;  
 DR EMBL; U47029; AAC50714.1; -;  
 DR EMBL; U47030; AAC50715.1; -;  
 DR EMBL; U47031; AAC50716.1; -;  
 DR EMBL; U47032; AAC50717.1; -;  
 DR EMBL; U47033; AAC50718.1; -;  
 DR EMBL; U47034; AAC50719.1; -;  
 DR EMBL; U47035; AAC50720.1; -;  
 DR EMBL; U47036; AAC50721.1; -;  
 DR EMBL; U47037; AAC50722.1; -;  
 DR EMBL; U47038; AAC50723.1; -;  
 DR EMBL; U47039; AAC50724.1; -;  
 DR EMBL; U47040; AAC50725.1; -;  
 DR EMBL; U47041; AAC50726.1; -;  
 DR EMBL; U47042; AAC50727.1; -;  
 DR EMBL; U47043; AAC50728.1; -;  
 DR EMBL; U47044; AAC50729.1; -;  
 DR EMBL; U47045; AAC50730.1; -;  
 DR EMBL; U47046; AAC50731.1; -;  
 DR EMBL; U47047; AAC50732.1; -;  
 DR EMBL; U47048; AAC50733.1; -;  
 DR EMBL; U47049; AAC50734.1; -;  
 DR EMBL; U47050; AAC50735.1; -;  
 DR EMBL; U47051; AAC50736.1; -;  
 DR EMBL; U47052; AAC50737.1; -;  
 DR EMBL; U47053; AAC50738.1; -;  
 DR EMBL; U47054; AAC50739.1; -;  
 DR EMBL; U47055; AAC50740.1; -;  
 DR EMBL; U47056; AAC50741.1; -;  
 DR EMBL; U47057; AAC50742.1; -;  
 DR EMBL; U47058; AAC50743.1; -;  
 DR EMBL; U47059; AAC50744.1; -;  
 DR EMBL; U47060; AAC50745.1; -;  
 DR EMBL; U47061; AAC50746.1; -;  
 DR EMBL; U47062; AAC50747.1; -;  
 DR EMBL; U47063; AAC50748.1; -;  
 DR EMBL; U47064; AAC50749.1; -;  
 DR EMBL; U47065; AAC50750.1; -;  
 DR EMBL; U47066; AAC50751.1; -;  
 DR EMBL; U47067; AAC50752.1; -;  
 DR EMBL; U47068; AAC50753.1; -;  
 DR EMBL; U47069; AAC50754.1; -;  
 DR EMBL; U47070; AAC50755.1; -;  
 DR EMBL; U47071; AAC50756.1; -;  
 DR EMBL; U47072; AAC50757.1; -;  
 DR EMBL; U47073; AAC50758.1; -;  
 DR EMBL; U47074; AAC50759.1; -;  
 DR EMBL; U47075; AAC50760.1; -;  
 DR EMBL; U47076; AAC50761.1; -;  
 DR EMBL; U47077; AAC50762.1; -;  
 DR EMBL; U47078; AAC50763.1; -;  
 DR EMBL; U47079; AAC50764.1; -;  
 DR EMBL; U47080; AAC50765.1; -;  
 DR EMBL; U47081; AAC50766.1; -;  
 DR EMBL; U47082; AAC50767.1; -;  
 DR EMBL; U47083; AAC50768.1; -;  
 DR EMBL; U47084; AAC50769.1; -;  
 DR EMBL; U47085; AAC50770.1; -;  
 DR EMBL; U47086; AAC50771.1; -;  
 DR EMBL; U47087; AAC50772.1; -;  
 DR EMBL; U47088; AAC50773.1; -;  
 DR EMBL; U47089; AAC50774.1; -;  
 DR EMBL; U47090; AAC50775.1; -;  
 DR EMBL; U47091; AAC50776.1; -;  
 DR EMBL; U47092; AAC50777.1; -;  
 DR EMBL; U47093; AAC50778.1; -;  
 DR EMBL; U47094; AAC50779.1; -;  
 DR EMBL; U47095; AAC50780.1; -;  
 DR EMBL; U47096; AAC50781.1; -;  
 DR EMBL; U47097; AAC50782.1; -;  
 DR EMBL; U47098; AAC50783.1; -;  
 DR EMBL; U47099; AAC50784.1; -;  
 DR EMBL; U47100; AAC50785.1; -;  
 DR EMBL; U47101; AAC50786.1; -;  
 DR EMBL; U47102; AAC50787.1; -;  
 DR EMBL; U47103; AAC50788.1; -;  
 DR EMBL; U47104; AAC50789.1; -;  
 DR EMBL; U47105; AAC50790.1; -;  
 DR EMBL; U47106; AAC50791.1; -;  
 DR EMBL; U47107; AAC50792.1; -;  
 DR EMBL; U47108; AAC50793.1; -;  
 DR EMBL; U47109; AAC50794.1; -;  
 DR EMBL; U47110; AAC50795.1; -;  
 DR EMBL; U47111; AAC50796.1; -;  
 DR EMBL; U47112; AAC50797.1; -;  
 DR EMBL; U47113; AAC50798.1; -;  
 DR EMBL; U47114; AAC50799.1; -;  
 DR EMBL; U47115; AAC50800.1; -;  
 DR EMBL; U47116; AAC50801.1; -;  
 DR EMBL; U47117; AAC50802.1; -;  
 DR EMBL; U47118; AAC50803.1; -;  
 DR EMBL; U47119; AAC50804.1; -;  
 DR EMBL; U47120; AAC50805.1; -;  
 DR EMBL; U47121; AAC50806.1; -;  
 DR EMBL; U47122; AAC50807.1; -;  
 DR EMBL; U47123; AAC50808.1; -;  
 DR EMBL; U47124; AAC50809.1; -;  
 DR EMBL; U47125; AAC50810.1; -;  
 DR EMBL; U47126; AAC50811.1; -;  
 DR EMBL; U47127; AAC50812.1; -;  
 DR EMBL; U47128; AAC50813.1; -;  
 DR EMBL; U47129; AAC50814.1; -;  
 DR EMBL; U47130; AAC50815.1; -;  
 DR EMBL; U47131; AAC50816.1; -;  
 DR EMBL; U47132; AAC50817.1; -;  
 DR EMBL; U47133; AAC50818.1; -;  
 DR EMBL; U47134; AAC50819.1; -;  
 DR EMBL; U47135; AAC50820.1; -;  
 DR EMBL; U47136; AAC50821.1; -;  
 DR EMBL; U47137; AAC50822.1; -;  
 DR EMBL; U47138; AAC50823.1; -;  
 DR EMBL; U47139; AAC50824.1; -;  
 DR EMBL; U47140; AAC50825.1; -;  
 DR EMBL; U47141; AAC50826.1; -;  
 DR EMBL; U47142; AAC50827.1; -;  
 DR EMBL; U47143; AAC50828.1; -;  
 DR EMBL; U47144; AAC50829.1; -;  
 DR EMBL; U47145; AAC50830.1; -;  
 DR EMBL; U47146; AAC50831.1; -;  
 DR EMBL; U47147; AAC50832.1; -;  
 DR EMBL; U47148; AAC50833.1; -;  
 DR EMBL; U47149; AAC50834.1; -;  
 DR EMBL; U47150; AAC50835.1; -;  
 DR EMBL; U47151; AAC50836.1; -;  
 DR EMBL; U47152; AAC50837.1; -;  
 DR EMBL; U47153; AAC50838.1; -;  
 DR EMB



```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95263670; PubMed=7744963;
RA Zhang H., Hu W., Ramirez F.;
RT "Developmental expression of fibrillin genes suggests heterogeneity
of extracellular microfibrils.";
RL J. Cell Biol. 129:1165-1176(1995).
RN [2]
RP SEQUENCE OF 210-317 FROM N.A.
RX MEDLINE=94140368; PubMed=8307578;
RA Li X., Pereira L., Zhang H., Sanguinetti C., Ramirez F., Bonadio J.,
RA Francke U.;
RT "Fibrillin genes map to regions of conserved mouse/human synteny on
mouse chromosomes 2 and 18.";
RL Genomics 18:667-672(1993).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L39790; AAA74908.1; -
DR EMBL; S69359; AAC60685.1; -
DR HSPSP; P35555; 1EMN.
DR MGD; MGI:95490; Fbn2.
DR InterPro; IPR000152; Asx hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002212; Fibril-assoc.
DR Pfam; PF00008; EGF; 46.
DR Pfam; PF00683; TB; 9.
DR PRINTS; PR00010; EGFLOOD.
DR SMART; SM00179; EGF_CA; 43.
DR SMART; SM00001; EGF_like; 3.
DR PROSITE; PS00010; ASX HYDROXYL; 43.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01185; EGF_2; 36.
DR PROSITE; PS01187; EGF_CA; 43.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 2907 FIBRILLIN 2.
FT DOMAIN 111 142 EGF-LIKE 1.
FT DOMAIN 145 176 EGF-LIKE 2.
FT DOMAIN 176 208 EGF-LIKE 3.
FT DOMAIN 276 317 EGF-LIKE 4, CALCIUM-BINDING.
FT DOMAIN 318 359 EGF-LIKE 5, CALCIUM-BINDING.
FT REPEAT 360 426 TGFBP 1.
FT DOMAIN 487 527 EGF-LIKE 6.
FT DOMAIN 528 567 EGF-LIKE 7, CALCIUM-BINDING.
FT DOMAIN 568 609 EGF-LIKE 8, CALCIUM-BINDING.
FT DOMAIN 610 650 EGF-LIKE 9, CALCIUM-BINDING.
FT DOMAIN 651 691 EGF-LIKE 10, CALCIUM-BINDING.
FT REPEAT 692 760 TGFBP 2.
FT DOMAIN 761 802 EGF-LIKE 11, CALCIUM-BINDING.
FT DOMAIN 803 844 EGF-LIKE 12, CALCIUM-BINDING.
FT DOMAIN 845 883 EGF-LIKE 13, CALCIUM-BINDING.
FT DOMAIN 948 989 EGF-LIKE 14, CALCIUM-BINDING.
FT REPEAT 990 1065 TGFBP 3.
FT DOMAIN 1066 1107 EGF-LIKE 15, CALCIUM-BINDING.
FT DOMAIN 1108 1150 EGF-LIKE 16, CALCIUM-BINDING.
FT DOMAIN 1151 1192 EGF-LIKE 17, CALCIUM-BINDING.
FT DOMAIN 1193 1234 EGF-LIKE 18, CALCIUM-BINDING.
FT DOMAIN 1235 1275 EGF-LIKE 19, CALCIUM-BINDING.
FT DOMAIN 1276 1317 EGF-LIKE 20, CALCIUM-BINDING.
FT DOMAIN 1318 1359 EGF-LIKE 21, CALCIUM-BINDING.
FT DOMAIN 1360 1400 EGF-LIKE 22, CALCIUM-BINDING.
FT DOMAIN 1401 1441 EGF-LIKE 23, CALCIUM-BINDING.
FT DOMAIN 1442 1483 EGF-LIKE 24, CALCIUM-BINDING.
FT DOMAIN 1484 1524 EGF-LIKE 25, CALCIUM-BINDING.
FT DOMAIN 1525 1565 EGF-LIKE 26, CALCIUM-BINDING.
FT REPEAT 1566 1642 TGFBP 4.
FT DOMAIN 1643 1684 EGF-LIKE 27, CALCIUM-BINDING.
FT DOMAIN 1685 1726 EGF-LIKE 28, CALCIUM-BINDING.
FT REPEAT 1727 1800 TGFBP 5.
FT DOMAIN 1801 1842 EGF-LIKE 29, CALCIUM-BINDING.
FT DOMAIN 1843 1884 EGF-LIKE 30, CALCIUM-BINDING.
FT DOMAIN 1885 1926 EGF-LIKE 31, CALCIUM-BINDING.
FT DOMAIN 1927 1965 EGF-LIKE 32, CALCIUM-BINDING.
FT DOMAIN 1966 2008 EGF-LIKE 33, CALCIUM-BINDING.
FT DOMAIN 2009 2048 EGF-LIKE 34, CALCIUM-BINDING.
FT DOMAIN 2049 2090 EGF-LIKE 35, CALCIUM-BINDING.
FT REPEAT 2091 2163 TGFBP 6.
FT DOMAIN 2164 2205 EGF-LIKE 36, CALCIUM-BINDING.
FT DOMAIN 2206 2245 EGF-LIKE 37, CALCIUM-BINDING.
FT DOMAIN 2246 2286 EGF-LIKE 38, CALCIUM-BINDING.
FT DOMAIN 2287 2330 EGF-LIKE 39, CALCIUM-BINDING.
FT DOMAIN 2331 2372 EGF-LIKE 40, CALCIUM-BINDING.
FT REPEAT 2373 2441 TGFBP 7.
FT DOMAIN 2442 2483 EGF-LIKE 41, CALCIUM-BINDING.
FT DOMAIN 2484 2524 EGF-LIKE 42, CALCIUM-BINDING.
FT DOMAIN 2525 2563 EGF-LIKE 43, CALCIUM-BINDING.
FT DOMAIN 2564 2606 EGF-LIKE 44, CALCIUM-BINDING.
FT DOMAIN 2607 2646 EGF-LIKE 45, CALCIUM-BINDING.
FT DOMAIN 2647 2687 EGF-LIKE 46, CALCIUM-BINDING.
FT DOMAIN 2688 2727 EGF-LIKE 47, CALCIUM-BINDING.
FT DISULFID 115 124 BY SIMILARITY.
FT DISULFID 119 130 BY SIMILARITY.
FT DISULFID 132 141 BY SIMILARITY.
FT DISULFID 149 159 BY SIMILARITY.
FT DISULFID 153 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 180 190 BY SIMILARITY.
FT DISULFID 184 196 BY SIMILARITY.
FT DISULFID 198 207 BY SIMILARITY.
FT DISULFID 280 292 BY SIMILARITY.
FT DISULFID 287 301 BY SIMILARITY.
FT DISULFID 303 316 BY SIMILARITY.
FT DISULFID 322 334 BY SIMILARITY.
FT DISULFID 329 343 BY SIMILARITY.
FT DISULFID 345 358 BY SIMILARITY.
FT DISULFID 491 503 BY SIMILARITY.
FT DISULFID 498 512 BY SIMILARITY.
FT DISULFID 514 526 BY SIMILARITY.
FT DISULFID 532 542 BY SIMILARITY.
FT DISULFID 537 551 BY SIMILARITY.
FT DISULFID 553 566 BY SIMILARITY.
FT DISULFID 572 584 BY SIMILARITY.
FT DISULFID 579 593 BY SIMILARITY.
FT DISULFID 595 608 BY SIMILARITY.
FT DISULFID 614 635 BY SIMILARITY.
FT DISULFID 620 634 BY SIMILARITY.
FT DISULFID 636 649 BY SIMILARITY.
FT DISULFID 655 666 BY SIMILARITY.
FT DISULFID 661 675 BY SIMILARITY.
FT DISULFID 677 690 BY SIMILARITY.
FT DISULFID 765 777 BY SIMILARITY.
FT DISULFID 772 786 BY SIMILARITY.
FT DISULFID 788 801 BY SIMILARITY.
FT DISULFID 807 819 BY SIMILARITY.
FT DISULFID 814 828 BY SIMILARITY.
FT DISULFID 830 843 BY SIMILARITY.
FT DISULFID 849 859 BY SIMILARITY.
FT DISULFID 854 868 BY SIMILARITY.
FT DISULFID 870 883 BY SIMILARITY.
FT DISULFID 952 964 BY SIMILARITY.

```







FT VARSPLIC 360 390 RLQVSQDIPSHLADEHALIASYVARLOHCAR -> S (IN ISOFORM 2).

FT VARSPLIC 526 558 AQATGSPHTSTHGGGPMMPVRSSTAGSTPT -> VSEL RNFPDSSRSFARSSFPFHINVSPLPAHHL (IN ISOFORM 3).

FT VARSPLIC 559 627 MISSING (IN ISOFORM 3).

FT VARSPLIC 579 608 MISSING (IN ISOFORM 2 AND ISOFORM 4).

SQ SEQUENCE 627 AA; 71355 MW; 6927EC0BD86D375C CRC64;

Alignment Scores:

Pred. No.:	6.95	Length:	627
Score:	77.00	Matches:	42
Percent Similarity:	31.41%	Conservative:	18
Best Local Similarity:	21.99%	Mismatches:	42
Query Match:	6.97%	Indels:	89
DB:	1	Gaps:	9

US-09-914-191-1 (1-598) x DTNB\_HUMAN (1-627)

QY 538 GACAAGATGAGATTGTTGTTTTCGAATTCACATCACTATAATAATGAGCTCCGATG 479

DB 146 AsplysLeuAArgTyrValPheSerGln-----Met 155

QY 478 TATAGGAAGGACTGCTTGCATATGGCAGAGGTGGATGAGTGCATGACAAAAGAAATT 419

DB 156 SerAspSerAsnGlyLeuMetilePheSerLysPheAspGlnPheLeuLysGluValLeu 175

QY 418 AAGCTGCCAACAGAAATG---GAAGGAAAAAAGATGGCAGTGCAGCCGCCAGCAGCAAG 362

DB 176 LysLeuProThrAlaValPheGluGlyProSerPheGlyTyrThr-----190

QY 361 CCAGTGCCTTGCACCTGGATATCATCGGGGATGCTTCTGGAAGGAACATCCAGAGATT 302

DB 191 -----GluHis-----192

QY 301 CTAGATGAAGACAGCTGACCTTTGCGCTTCAGTTCGTGTCGCTTA-----254

DB 193 -----SerValArgThrCysPheProGlnGlnArg 202

QY 253 -----ACCATGAAGCCCTCCACCTCCAGCGGCTCT 221

DB 203 LysileMetLeuAsnMetPheLeuAspThrMetMetAlaAspPropProGln-----220

QY 220 TGCCTTAGTGGCTG-----TAGCATCCCTACCAACCA 188

DB 221 CysLeuValTrpLeuProLeuMetHisArgLeuAlaHisValGluAsnValPheHisPro 240

QY 187 GGACAC-----TGGTGGGAATGACACAACTCAAGTTGGGAGGGAACAGGAAGG 134

DB 241 ValGluCysSerTyrCysArg-----247

QY 133 GATGGATGGGGTGGTGTATCTTACTCTGTTTAAGCAGAACACCTTGTGTCGGTGTGG 74

DB 248 -----CysGluSerMetMetGlyPheArgTyrArgCysGlnGlnCysHis 262

QY 73 AACATGGTTCCTTTGGCAGAGTGCCTTTT 43

DB 263 Asn-TyrGlnLeuCysGlnAsnCysPheTrp 272

## RESULT 13

GAL\_PSEFL ID - GAL\_PSEFL STANDARD; PRT; 304 AA.

AC P11886;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 01-DEC-1992 (Rel. 24, Last annotation update)

DE D-galactose 1-dehydrogenase (EC 1.1.1.48).

GN GAL.

OS Pseudomonas fluorescens.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI\_taxID=294;

RN (1)

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-40.

RC STRAIN=BMTU 102;

RX MEDLINE=89345117; PubMed=2503815;

RA Sperka S., Zehelein E., Fiedler S., Fischer S., Sommer R., Buckel P.;

RT "Complete nucleotide sequence of Pseudomonas fluorescens D-galactose dehydrogenase gene.";

RL Nucleic Acids Res. 17:5402-5402(1989).

CC -1- CATALYTIC ACTIVITY: D-galactose + NAD(+) = D-galactono-1,4-lactone + NADH.

CC -1- PATHWAY: FIRST STEP IN D-GALACTOSE CATABOLISM.

CC -1- SUBUNIT: HOMODIMER.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X14913; CAA33038.1; -

DR PIR; S04853; S04853

DR InterPro; IPR000683; GFO\_IDH\_MOCA.

DR InterPro; IPR004104; GFO\_IDH\_MOCA\_C.

DR Pfam; PF01408; GFO\_IDH\_MOCA; 1.

DR Pfam; PF02894; GFO\_IDH\_MOCA\_C; 1.

KW Oxidoreductase; NAD; Galactose metabolism.

SQ SEQUENCE 304 AA; 33049 MW; D842A776DC67A3B0 CRC64;

Alignment Scores:

Pred. No.:	7.88	Length:	304
Score:	76.00 <td>Matches:</td> <td>35 </td>	Matches:	35
Percent Similarity:	39.43% <td>Conservative:</td> <td>34 </td>	Conservative:	34
Best Local Similarity:	20.00% <td>Mismatches:</td> <td>58 </td>	Mismatches:	58
Query Match:	6.93% <td>Indels:</td> <td>48 </td>	Indels:	48
DB:	1 <td>Gaps:</td> <td>6 </td>	Gaps:	6

US-09-914-191-1 (1-598) x GAL\_PSEFL (1-304)

QY 10 TTCTTCTTTTATAAAATAGTACTGCGATTAAAAAAGACACTTCTGCCAAGGAACCA 69

DB 73 PheAlaLeuValGlnGlnAlaLeuAlaGlyLysHisValLeuValGluLysProPro 92

QY 70 TGTTCACACACCCGCAACCAAGGTGTTCTGTTAAACAGAGTAAGATACACACCCCATC 129

DB 93 CysAlaThrLeuGlyLysAlaAlaLeuTrpLysArgGluGlnAlaSerAlaProCys 112

QY 130 CATCCCTTCCTCCCTGTTCCCTCCCACTTGAGTTGTGTGTCATTCGCACAGTGTCTG 189

DB 113 SerProCysileAlaTyrAlaPro-----AlaIleAla 123

QY 190 GGTGGTAGGATGCTACAGCCACCTAAGCAAGGAGCCCTGGGAGGTGGAGGGCTGCA 249

DB 124 AlaAlaArgAsp-----127

QY 250 TGTGTAAGCACACACAGAACTGAAGCGCAAAAGGTCAGCTGCTTCTCATCTAGAATCTCTG 309

DB 128 TrpLeuAlaThrArgThr-----LeuGlnSerVal 137

QY 310 GATGTTCTTCCAGAAAGATCCCGATGATATCGCAGTGCAGGACGCTGGCTTGTGTC 369

DB 138 GlnIleAspTrpLysGluAspValArgLysTrpHisProGlyGlnAla-TripIleTrpG 157

QY 370 TGTGCGGGTCACTGCCATCTTTTCTTCATTTCTGTTGGCAGCTTAATTTCTTTG 429

DB 157 n---ProGly-LeuGlyValPheAspProGlyIleAsnAlaLeuSerIleValThr---174

QY 430 TCATCACTTCATCCACCTTCTGCATATCAACACAGTCCCTTTCTTATATACATCGCAG 489

DB 175 -----HisLeuLeuProLeuPro---LeuPheValGluSerA 186

QY 490 CATTATTATAGTTGATGTTGAATTCAGAAACAAAATCTCA 530

```
Db 186 laGluLeuArgValProSerAsnCysGlnSerProIleAla 199
:: ||| ::| ||| ::| ::| ::| ::| ::| ::| ::| ::|
AC 070624; 070289; STANDARD; PRT; 490 AA.
ID MYOC MOUSE
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myocilin precursor (Trabecular meshwork-induced glucocorticoid
DE response protein).
GN MYOC OR TIGR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=98217378; PubMed=9548973;
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=129/SV;
RX MEDLINE=98345432; PubMed=9680392;
RA Abderrahim H., Jaramillo-Babb V.L., Zhou Z., Vollrath D.;
RT "Characterization of the murine TIGR/myocilin gene.";
RL Mamm. Genome 9:673-675(1998).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98217378; PubMed=9548973;
RA Fingert J.H., Ying L., Swiderski R.E., Nystuen A.M., Arbour N.C.,
RA Alward W.D.M., Sheffield V.C., Stone E.M.;
RT "Characterization and comparison of the human and mouse GLC1A glaucoma
RL genes.";
RL Genome Res. 8:377-384(1998).
[3]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND VARIANT ALA-164.
RC STRAIN=BALB/cJ, C3H/HeJ, and C57BL/6J; TISSUE=Brain, and Muscle;
RX MEDLINE=98249809; PubMed=9588210;
RA Tomarev S.I., Tamm E.R., Chang B.;
RT "Characterization of the mouse Myoc/Tigr gene.";
RL Biochem. Biophys. Res. Commun. 245:887-893(1998).
[4]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=BALB/c; TISSUE=skeletal muscle;
RX MEDLINE=98340858; PubMed=9675094;
RA Takahashi H., Noda S., Imamura Y., Nagasawa A., Kubota R., Mashima Y.,
RA Kudoh J., Oguchi Y., Shimizu N.;
RT "Mouse myocilin (Myoc) gene expression in ocular tissues.";
RL Biochem. Biophys. Res. Commun. 248:104-109(1998).
CC -!- FUNCTION: MAY PARTICIPATE IN THE OBSTRUCTION OF FLUID OUTFLOW IN
CC THE TRABECULAR MESHWORK.
CC -!- SUBCELLULAR LOCATION: LOCATED PREFERENTIALLY IN THE CILIARY
CC ROOTLET AND BASAL BODY OF THE CONNECTING CILIUM OF PHOTORECEPTOR
CC CELLS, AND IN THE ROUGH ENDOPLASMIC RETICULUM. ALSO SECRETED.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN CILIARY BODY, IRIS, RETINA,
CC TRABECULAR NETWORK AND SCLERA BUT NOT IN LENS OR CORNEA. ALSO
CC EXPRESSED STRONGLY IN SKELETAL MUSCLE AND WEAKLY IN HEART, BRAIN,
CC TESTIS, LIVER, KIDNEY, THYROID AND EPIDIDYMIS. NO EXPRESSION
CC DETECTED IN EMBRYO.
CC -!- POLYMORPHISM: VARIANT ALA-164 IS FOUND IN STRAIN BALB/cJ WHICH HAS
CC A LOW INTRAOCULAR PRESSURE. VARIANT THR-164 IS FOUND IN STRAINS
CC C3H/HeJ AND C57BL/6J, TWO STRAINS WHICH HAVE A RELATIVELY HIGH
CC INTRAOCULAR PRESSURE.
CC -!- SIMILARITY: BELONGS TO THE OLFACOMEDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF041335; AAC32805.1; -.
DR EMBL; AF041333; AAC32805.1; JOINED.
```

```
DR EMBL; AF041334; AAC32805.1; JOINED.
DR EMBL; AF049796; AAC14265.1; -.
DR EMBL; AF049795; AAC14265.1; JOINED.
DR EMBL; AF049794; AAC14265.1; JOINED.
DR EMBL; AF039869; AAC40112.1; -.
DR EMBL; AB013592; BAA32031.1; -.
DR MGD; MGI:1202864; MYOC.
DR InterPro; IPR003112; Olfac_like.
DR Pfam; PF02191; OLF; 1.
DR SMART; SM00284; OLF; 1.
KW Coiled coil; Glycoprotein; Signal; Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 490
FT DOMAIN 69 170
FT DOMAIN 232 490
FT CARBOHYD 43 43
FT VARIANT 164 164
SQ SEQUENCE 490 AA; 55314 MW; 2F090571E97B0425 CRC64;

Alignment Scores:
Pred. No.: 8.49 Length: 490
Score: 76.00 Matches: 49
Percent Similarity: 33.87% Conservative: 14
Best Local Similarity: 26.34% Mismatches: 60
Query Match: 6.88% Indels: 63
DB: 1 Gaps: 10

US-09-914-191-1 (1-598) x MYOC_MOUSE (1-490)
QY 574 CAAGCCCAAGGAGATTACAGGAACTCTTNCAGCA-----GACAAAGAT 530
Db 98 GluAlaGlnGluGlyLeuGlnGlyAlaLeuArgGluArgGlnLeu 117
QY 529 GAGATTTTCTTGAATTCAACTAATCACTAATAATAGCTGCCGATGATAGAAA 470
Db 118 GluThrGlnThrArgAspLeuGluAlaIatyrAsnAsnLeuArg----- 133
QY 469 GGGACTGTGTGATATGGCAGAGGTGATGAAGTGCACAAAGAAATAAGCTGCCA 410
Db 134 -----AspLysSerAlaLeuGluGluGluGln 144
QY 409 ACAGAAATGGAAGGAAAAAGATGGCA-----GTGACC 377
Db 145 LeuGluGlnGluAsnGluAspLeuAlaArgLeuGluSerSerGluGluValThr 164
QY 376 CGGACACAGACAAAGCCAGTCGCTGCGACTGGATA---TCATCGGGATGCTTCTGG 320
Db 165 ArgLeuArg-ArgGlyGlnCysProSerThrGlnTyProSerGlnAspMetLeuProG 184
QY 319 A---AGCAACATCCAGAGATTCTAGATGAAGACAGCTGACCCCTTTTCGCTTCAGTCTG 263
Db 184 yserArgGluValSerGlnTrp-----AsnLeuAspThrLeuAlaPheGln----- 199
QY 262 GTGTGCTTAACCATGCAAGCCCTCCACCTCCAGGGCTCTTTCCTTAGGTGGCTGTAG 203
Db 200 -----GluLeuLysSerGluLeuThr-----G1 207
QY 202 CATCCCTACACCCAGGACACTGGTGGATGACAACTCAAGTTGGGAGGGAACAGG 143
Db 207 uValProAlaSerGlnIleLeuLysGluAsnProSer-----GlyArgProArgse 224
QY 142 GAAGGAAGGATGGATGGGGGTGTGTATCTTACTCTGTTTAAGCAGAACACCTCTGTTG 83
Db 224 rlysgluGlyAspLysGly-----Cyt 231
QY 82 CGGTGTGGAACATGG 67
Db 231 sGlyAlaLeuValTrp 236

RESULT 15
P3K1_DICDI
ID P3K1_DICDI
AC P54673;
PRT; 1570 AA.
```

DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Phosphatidylinositol 3-kinase 1 (EC 2.7.1.137) (PI3-Kinase)  
DE (PtdIns-3-kinase) (PI3K).  
GN PI3K OR PIK1.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AX3.  
RX MEDLINE=96009592; PubMed=7565716;  
RA Zhou K., Takegawa K., Emr S.D., Firtel R.A.;  
RT "A phosphatidylinositol (PI) kinase gene family in Dictyostelium  
discoidium: biological roles of putative mammalian p110 and yeast  
Vps34p PI 3-kinase homologs during growth and development.";  
RL Mol. Cell. Biol. 15:5645-5656(1995).  
CC -!- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-ID-myo-inositol = ADP +  
1-phosphatidyl-ID-myo-inositol 3-phosphate.  
CC -!- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U23476; AAA85721.1; -.  
DR DictyDb; DD01099; pikA.  
DR InterPro; IPR002420; PI3K\_C2.  
DR InterPro; IPR000341; PI3K\_ras\_bind.  
DR InterPro; IPR001263; PI3Ka.  
DR InterPro; IPR000403; PI3\_Pi4\_kinase.  
DR InterPro; IPR000626; Ubiquitin.  
DR Pfam; PF00240; ubiquitin; 1.  
DR Pfam; PF00454; PI3\_Pi4\_kinase; 1.  
DR Pfam; PF00613; PI3Ka; 1.  
DR Pfam; PF00792; PI3K\_C2; 1.  
DR Pfam; PF00794; PI3K\_rbd; 1.  
DR SMART; SM00142; PI3K\_C2; 1.  
DR SMART; SM00144; PI3K\_rbd; 1.  
DR SMART; SM00145; PI3Ka; 1.  
DR SMART; SM00146; PI3K; 1.  
DR PROSITE; PS00915; PI3\_4\_KINASE\_1; 1.  
DR PROSITE; PS00916; PI3\_4\_KINASE\_2; 1.  
DR PROSITE; PS0290; PI3\_4\_KINASE\_3; 1.  
KW Transferase; Kinase; Multigene family.  
FT DOMAIN 41 48 POLY-ASN.  
FT DOMAIN 59 73 POLY-ASN.  
FT DOMAIN 161 168 POLY-ASN.  
FT DOMAIN 294 304 POLY-THR.  
FT DOMAIN 308 324 POLY-THR.  
FT DOMAIN 413 424 POLY-ASN.  
FT DOMAIN 503 510 POLY-SER.  
FT DOMAIN 570 579 POLY-ASN.  
FT DOMAIN 821 828 POLY-ASN.  
FT DOMAIN 831 836 POLY-ASN.  
FT DOMAIN 1309 1570 PI3K/PI4K.  
SQ SEQUENCE 1570 AA; 178374 MW; 55B678B72B34D783 CRC64;

## Alignment Scores:

Pred. No.:	10.2	Length:	1570
Score:	76.00	Matches:	32
Percent Similarity:	35.20%	Conservative:	12
Best Local Similarity:	25.60%	Mismatches:	27
Query Match:	6.88%	Indels:	54
DB:	1	Gaps:	7

US-09-914-191-1 (1-598) x P3K1\_DICD1 (1-1570)

QY	508	AACATCACTAATAATAGAGTCCGATCTATAGGAAGGGACTGTGTGATATATGCAG	449
Db	834	AsnIleAsnPheAspAsn-----	841
QY	448	AAGGTGGATGAAGTATGACAAAAGAAATTAAGCTGCCAACAGAAATGGAAGGAAAAG	389
Db	842	GlnIleThrHisIleSerIleA:GGluIleLysProPheArgVal-----LysVal	859
QY	388	ATGGCAGTGAACCGGACCCAGCAAAAGCCAGTGCCTTGCACCTGCGATATCATCGGGGAT	329
Db	860	MetGlySerThrArg-----IleProLeuSerCys-----	869
QY	328	GCTTTCTGGAAGGAACATCCAGAGATCTAGATGAAGACAGACGACCTTTTGGCGTTCA	269
Db	870	-----IleLysAspIleAspSer-----SerSer	877
QY	268	GTTCCTGGTGTGCTTA-----	248
Db	878	ValIleValSerIleSerLeuTyrHisGlyIleGluCysPheSerLysAlaPheThrGln	897
QY	247	CAAGCCCTCCACCTCCCGGGCTCCTTGCCTTAGGTGGCTGTAGCATCCCTACCACCCA	188
Db	898	ProIleIleProProPheAlaPheLeuAlaGluThrLeu-----Ser	912
QY	187	GGACACTGGTGGCAA	173
Db	913	ValAspTrpCysGlu	917

Search completed: July 9, 2003, 12:48:48  
Job time : 22 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: July 9, 2003, 12:45:58 ; Search time 47.5 Seconds

(without alignments)

5188.049 Million cell updates/sec

Title: US-09-914-191-1

Perfect score: 1096

Sequence: 1 ttggaatagttcttctgctta.....gggttagtcagattgttg 598

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p model -DEV=xlp  
-Q=/cgn2.1/USPTO\_spool/US0914191/runat\_09072003.112349.12876/app\_query.fasta.1.775  
-DB=SPTRMBL\_21 -Qfmt=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=spct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US0914191@cgn 1.1.138 @runat\_09072003.112349.12876 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	544	49.2	173	4	Q9H8R6

C	2	544	49.2	298	4	Q9H0S2
C	3	537	48.6	298	4	Q9NWX6
C	4	428	38.7	238	11	Q9CQT0
C	5	428	38.7	298	11	Q9CY52
C	6	219	19.8	286	5	Q9V3N8
C	7	169.5	15.3	261	3	Q9Y7T3
C	8	129.5	11.7	495	10	Q9SIQ0
C	9	94	8.5	297	10	Q9ZV61
C	10	88.5	8.1	457	10	Q93ZL5
C	11	88.5	8.1	515	10	Q9FK95
C	12	85	7.7	330	12	Q85277
C	13	84.5	7.7	495	11	Q9Z1R1
C	14	82.5	7.5	343	10	Q9LP79
C	15	82	7.5	1628	4	Q9HCL7
C	16	81.5	7.4	370	10	Q22742
C	17	81.5	7.4	661	10	Q9FVC8
C	18	81.5	7.4	2406	4	Q9B2S0
C	19	81	7.3	390	12	Q85276
C	20	80.5	7.3	410	16	Q8YSZ6
C	21	80	7.2	304	3	Q9C1J4
C	22	80	7.3	402	4	Q9NS06
C	23	80	7.3	403	4	Q9C009
C	24	80	7.3	484	4	Q9B204
C	25	79.5	7.3	2144	11	Q9Q1P2
C	26	79	7.2	359	10	Q9AVE6
C	27	78.5	7.2	225	5	Q9NAF1
C	28	78.5	7.2	397	16	Q9K7I5
C	29	78	7.1	252	17	Q8TZ46
C	30	78	7.1	672	11	Q9WVH4
C	31	78	7.1	2906	11	Q9WU9H
C	32	77	7.0	181	4	Q9UPC2
C	33	77	7.0	560	4	Q96AW0
C	34	77	7.0	683	11	Q9DSW6
C	35	76.5	6.9	264	16	Q84478
C	36	76.5	6.9	284	16	Q9PUS1
C	37	76.5	7.0	293	3	Q12393
C	38	76.5	6.9	439	2	Q60021
C	39	76.5	6.9	612	10	Q9FGW5
C	40	76.5	7.0	644	15	Q85735
C	41	76.5	6.9	1006	3	Q9C2A4
C	42	76.5	6.9	1037	16	Q9KRG9
C	43	76.5	7.0	1773	10	Q94HV6
C	44	76	6.9	268	4	Q9UJD7
C	45	76	6.9	327	16	Q8Z0H4

#### ALIGNMENTS

##### RESULT 1

Q9H8R6 PRELIMINARY; PRT; 173 AA.  
ID Q9H8R6  
AC Q9H8R6;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE CDNA FLJ13293 fis, clone OVARC1001188 (Similar to hypothetical protein FLJ20546) (Hypothetical 20.2 kDa protein).  
DE Homo sapiens (Human)  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=OVARY;  
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosoi T., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
RA Ninomiya K., Iwayanagi T.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

RN SEQUENCE FROM N.A.  
 RP TISSUE=MUSCLE;  
 RC Strausberg R.;  
 RA Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=MUSCLE;  
 RA Strausberg R.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK023355; BAB14540.1; -;  
 DR EMBL; BC001852; AAH01852.1; -;  
 DR EMBL; BC001523; AAH01523.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 173 AA; 20157 MW; 70C1D8FFC116B416 CRC64;

Alignment Scores:  
 Pred. No.: 3,6e-52 Length: 173  
 Score: 544.00 Matches: 103  
 Percent Similarity: 99.04% Conservatives: 0  
 Best Local Similarity: 99.04% Mismatches: 1  
 Query Match: 49.23% Indels: 0  
 DB: 4 Gaps: 0

US-09-914-191-1 (1-598) x Q9H8R6 (1-173)

QY 598 CAACAATCTGGACTTAACACCACTACAGCCCAAGGAGATTACAGGAACTCTTNCAGCA 539  
 |||||  
 Db 70 GlnGlnSerGlyLeuThrProValGlnAlaGlnGlyArgLeuGlnGlyThrLeuAlaA 89  
 |||||

QY 538 GACAAGAATGAGATTTGTTCTTGAATTCACCACTCAACTATATAATGAGTCCCGATG 479  
 |||||  
 Db 90 AspLysAsnGluLeuPheSerGluPheAsnIleAsnTyrAsnAsnGluLeuPromet 109

QY 478 TATAGAAAGGAGCTGTGTTGATATATGCGAAGAGTGGATGATGATGATGATGATGAT 419  
 |||||  
 Db 110 TyrArgLysGlyThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 129

QY 418 AAGCTGCCAACAGAAATGGAGGAAAGAAAGATGCGAGTGGATGATGATGATGATGATGAT 359  
 |||||  
 Db 130 LysLeuProThrGluMetGluGlyLysValMetAlaValThrArgThrArgThrLysPro 149

QY 358 GTGCCCTTGCATCGCATATCATCGGGATGCTTTCTTGAAGGAAACATCCAGAGATTCTA 299  
 |||||  
 Db 150 ValProLeuHisCysAspIlelleGlyAspAlaPheThrLysGluHisProGluLeuLeu 169

QY 298 GATGAAGACAGC 287  
 |||||  
 Db 170 AspGluAspSer 173

RESULT 2  
 Q9H8S2 ID Q9H8S2 PRELIMINARY; PRT; 298 AA.  
 AC Q9H8S2;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hypothetical 34.8 kDa protein.  
 GN DKF2P564B1172.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRIN;  
 RX MEDLINE=21154917; PubMed=11230166;  
 RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,  
 RA Ansoerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
 RA Lauber J., Duesterhoeft A., Beyer A., Koehner K., Strack N.,  
 RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,  
 RA Wambutt R., Korn B., Klein M., Poustka A.;  
 \*RT "Towards a Catalog of Human Genes and Proteins: Sequencing and

RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";  
 RL Genome Res. 11:422-435 (2001).  
 DR EMBL; AL136669; CAB66604.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 298 AA; 34831 MW; A6C941B75611C448 CRC64;

Alignment Scores:  
 Pred. No.: 3,84e-52 Length: 298  
 Score: 544.00 Matches: 103  
 Percent Similarity: 99.04% Conservatives: 0  
 Best Local Similarity: 99.04% Mismatches: 1  
 Query Match: 49.23% Indels: 0  
 DB: 4 Gaps: 0

US-09-914-191-1 (1-598) x Q9HOS2 (1-298)

QY 598 CAACAATCTGGACTTAACACCACTACAGCCCAAGGAGATTACAGGAACTCTTNCAGCA 539  
 |||||  
 Db 195 GlnGlnSerGlyLeuThrProValGlnAlaGlnGlyArgLeuGlnGlyThrLeuAlaA 214  
 |||||

QY 538 GACAAGAATGAGATTTGTTCTTGAATTCACCACTCAACTATATAATGAGTCCCGATG 479  
 |||||  
 Db 215 AspLysAsnGluLeuPheSerGluPheAsnIleAsnTyrAsnAsnGluLeuPromet 234

QY 478 TATAGAAAGGAGCTGTGTTGATATATGCGAAGAGTGGATGATGATGATGATGATGAT 419  
 |||||  
 Db 235 TyrArgLysGlyThrValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 254

QY 418 AAGCTGCCAACAGAAATGGAGGAAAGAAAGATGCGAGTGGATGATGATGATGATGATGAT 359  
 |||||  
 Db 255 LysLeuProThrGluMetGluGlyLysValMetAlaValThrArgThrArgThrLysPro 274

QY 358 GTGCCCTTGCATCGCATATCATCGGGATGCTTTCTTGAAGGAAACATCCAGAGATTCTA 299  
 |||||  
 Db 275 ValProLeuHisCysAspIlelleGlyAspAlaPheThrLysGluHisProGluLeuLeu 294

QY 298 GATGAAGACAGC 287  
 |||||  
 Db 295 AspGluAspSer 298

RESULT 3  
 Q9NWX6 ID Q9NWX6 PRELIMINARY; PRT; 298 AA.  
 AC Q9NWX6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TrEMBLrel. 18, Last annotation update)  
 DE Hypothetical 34.8 kDa protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
 RA Suzuki Y., Obayashi M., Nishi T., Shibahara T., Tanaka T.,  
 RA Nakamura Y., Isogai T., Sugano S.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=EMBRYO;  
 RA Isogai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Wagatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,  
 RA Ninomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK000553; BAA91249.1; -;  
 DR EMBL; AK021663; BAB13870.1; -;



RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK010876; BAB27240.1; -;  
 DR MGI; MGI:1913878; 573040907Rik.  
 SQ SEQUENCE 298 AA; 34971 MW; D4CA0C154E452ECA CRC64;  
 Alignment Scores:  
 Pred. No.: 4.06e-39 Length: 298  
 Score: 428.00 Matches: 80  
 Percent Similarity: 89.42% Conservatives: 13  
 Best Local Similarity: 76.92% Mismatches: 11  
 Query Match: 38.73% Indels: 0  
 DB: 11 Gaps: 0  
 US-09-914-191-1 (1-598) x 09CY52 (1-298)  
 QY 598 CAACAATCTGGACTTAACACCAAGCCAGGAGATACAGGAACTCTTNCAGCA 539  
 Db 195 GlnGlnSerGlyLeuThrProValGlnAlaGlnGlnArgLeuGlyThrLeuThrAla 214  
 QY 538 GACAGAAATGAGATTTTCTTCTGATTCACATCACTATATATGAGCTGCCGATG 479  
 Db 215 AspLysAsnGluLeuLeuPheSerGluPheHisIleAsnTyrAsnAsnGluProHisMet 234  
 QY 478 TATAGGAAGGAGCTGTGTGATATGCGACAGAGTGATGATGACAAAGAAAT 419  
 Db 235 TyrArgLysGlyThrValLeuValTrpGlnLysValGluValArgThrGlnGluVal 254  
 QY 418 AAGCTGCCACAGAAATGGAAGGAAAGATGCGAGTACCGCGACCGACCAAGCCCA 359  
 Db 255 ArgLeuProAlaGluMetGluGlyGlyLysAlaValAlaArgThrArgThrArgVal 274  
 QY 358 GTGCCCTTGCACTGCGATATCATCGGGATCGCTTCTGGAAGCAATCAGAGATCTA 299  
 Db 275 ValAlaLeuAsnCysAspLeuLeuGlyAspAlaPheTrpLysGluHisProGluLeu 294  
 QY 298 GATGACACAGC 287  
 Db 295 AlaGluGluAsn 298  
 RESULT 6  
 ID Q9V3N8 PRELIMINARY; PRT; 286 AA.  
 AC Q9V3N8;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 13, Last sequence update)  
 DE BG:DS00929.4 protein.  
 GN BG:DS00929.4 OR CG4103.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Lian X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=99403001; PubMed=10471707;  
 RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,  
 RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,  
 RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,  
 RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,  
 RA Celniker S., Rubin G.M.;  
 RT "An exploration of the sequence of a 2.9-Mb region of the genome of  
 RT Drosophila melanogaster: the Adh region.";  
 RL Genetics 153:179-219(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Celniker S.E., Agbayani A., Arcaina T.T., Baxter E., Blazej R.G.,  
 RA Butanoff C., Champe M., Chavez C., Chew M., Ciesiolka L., Doyle C.M.,  
 RA Farfan D.E., Galle R., George R.A., Harris N.L., Hoskins R.A.,  
 RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,  
 RA Lewis S., Li P., Lomotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,  
 RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,  
 RA Sethi H., Shih E., Svirskas R.R., Wan K.H., Weinburg T., Zhang R.,  
 RA Zieran L.L., Rubin G.M.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003646; AAF53429.1; -;  
 DR EMBL; AE003411; AAF44908.1; -;  
 DR FlyBase; FBgn0028925; BG:DS00929.4.  
 KW Hypothetical protein.  
 SQ SEQUENCE 286 AA; 33072 MW; 9539693DA9C2BFC8 CRC64;  
 Alignment Scores:  
 Pred. No.: 1.18e-15 Length: 286  
 Score: 219.00 Matches: 47  
 Percent Similarity: 60.61% Conservatives: 13  
 Best Local Similarity: 47.47% Mismatches: 21  
 Query Match: 19.82% Indels: 18  
 DB: 5 Gaps: 2  
 US-09-914-191-1 (1-598) x Q9V3N8 (1-286)  
 QY 595 CAATCTGGACTTAACACCAAGCCAGGAGATACAGGAACTCTTNCAGCAGAC 536  
 Db 166 GluLysGlyLeuThrAsnGlnGlnAlaGluLysLeuArgGlyThrPheSerAlaAsp 185  
 QY 535 AGAATGAGATTTTCTTCTGATTCACATCACTATATATGAGCTGCCGATGAT 476  
 Db 186 LysAsnGluLeuLeuPheGlnGluPheGlyIleAsnTyrAsnAsnLeuProAlaMetTyr 205  
 QY 475 AGGAAAGGAGCTGTGTGATGATGCGAGAGGTGGTGAAGTATGATGACAAAAGAAATTAAG 416



```

||||| 206 ArgLysGlyThrIleLeuLeuArgLysArgVal 216
QY 415 CTGCCAACAGAAATGGAAGCAAAAGATGGCAGTGCCCGAGCAGGACCAAGCCAGT 356
DB 217 -----IleLeuGlyGluLys-----SerArgGlnAlaVal 227
QY 355 CCCTTGCACTCGATATCATCGGGATCTTCTGGAAGGAACATCCAGAGATTCTA 299
DB 228 ProLeuHisGluAspLeuIleSerSerGlnPheTrpLysGluHisThrGluIleLeu 246

RESULT 7
Q9Y7T3 PRELIMINARY; PRT; 261 AA.
AC Q9Y7T3;
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE Hypothetical 31.1 kDa protein C63.07 in chromosome III.
GN SPC63.07.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -i- SIMILARITY: TO YEAST YGR024C.
DR EMBL; AL049522; CAB40011.1; -.
KW Hypothetical protein.
SQ SEQUENCE 261 AA; 31063 MW; FCC4B047C3D2D7DA CRC64;

```

```

Alignment Scores:
Pred. No.: 4,22e-10 Length: 261
Score: 169.50 Matches: 38
Percent Similarity: 59.41% Conservative: 22
Best Local Similarity: 37.62% Mismatches: 32
Query Match: 15.34% Indels: 9
DB: 3 Gaps: 2

```

```

US-09-914-191-1 (1-598) x Q9Y7T3 (1-261)
QY 595 CAATCTGGACTAACACAGTACAGCCAGGAGATTACAGGAACTCTTNCAGCAGAC 536
DB 166 LysGlyGlyPheThrAsnThrGlnAlaGluGluTyrLeuLysGlyThrValSerAlaGlu 185
QY 535 AAGAAAGAGATTTGTTTCTGAATTCACATCAACTATAATAGCTGCCGATGTAT 476
DB 186 LysHisGluIleLeuPheSerLysPheGlyIleAsnTyrAsnPheGluProGluIleTyr 205
QY 475 AGAAAGGAGCTGTGTTGATATGCGAGAGTGGATGAAGTATGACAAAGAAATTAAG 416
DB 206 LysLysGlySerIleTrpIleArgGluProIleAspGlnGluTrpHisGlnAspLys 225
QY 415 CTGCCAACAGAAATGGAAGCAAAAGATGCGAGTACCCCGAGCAGGACCAAGCCAGT 356
DB 226 LysPheSerValLysGlnLysLysLysMetValLeuSer----- 238
QY 355 CCCTTGCACTCGATATCATCGGGATCTTCTGGAAGGAACATCCA-----GAGATT 302
DB 239 IleLeuHisValSerLeuIleAspAspPheTrpThrSerArgPropheLeuGluVal 258
QY 301 CTA 299
DB 259 Leu 259

```

```

RESULT 8
Q9SIQ0 PRELIMINARY; PRT; 495 AA.
ID Q9SIQ0
AC Q9SIQ0;

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 57.9 kDa protein.
GN Ar2G31580.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007071; AAD24854.2; -.
KW Hypothetical protein.
SQ SEQUENCE 495 AA; 57942 MW; E75CABDB892453BE CRC64;

```

```

Alignment Scores:
Pred. No.: 1.41e-05 Length: 495
Score: 129.50 Matches: 32
Percent Similarity: 51.96% Conservative: 21
Best Local Similarity: 31.37% Mismatches: 40
Query Match: 11.72% Indels: 9
DB: 10 Gaps: 2

```

US-09-914-191-1 (1-598) x Q9SIQ0 (1-495)

```

QY 595 CAATCTGGACTAACACAGTACAGCCAGGAGATTACAGGAACTCTTNCAGCAGAC 536
DB 166 LysSerGlyLysSerValSerGluThrGlnGluIleLeuLysAspThrGlnLysGlnGln 185
QY 535 AAGAAAGAGATTTGTTTCTGAATTCACATCAACTATAATAGCTGCCGATGTAT 476
DB 186 LysAsnGluLeuPheGlnLysPheGlyIleAsnTyrLysThrLeuProGluLeuPhe 205
QY 475 AGAAAGGAGCTGTGTTGATATGCGAGAGTGGATGAAGTATGACAAAGAAATTAAG 416
DB 206 ArgGlnGlySerCysLeuPheLysLysLysValGluGluThrVal----- 220
QY 415 CTGCCAACAGAAATGGAAGCAAAAGATGCGAGTACCCCGAGCAGGACCAAGCCAGT 356
DB 221 -----LysHisaspGluAsnGlyAsnProValLysArgLeuArgLysAlaVal 237
QY 355 CCCTTGCACTCGCAT---ATCATCGGGATGCTTCTGGAAGGAACATCCAGAGATTCTA 299
DB 238 PheValHisSerGluAsnIleAlaGlyArgSerPheTrpAsnGluGlnProSerLeuTyr 257
QY 298 GATGAA 293
DB 258 AsnAsp 259

```

RESULT 9

```

Q9ZV61 PRELIMINARY; PRT; 297 AA.
ID Q9ZV61
AC Q9ZV61;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE At2G32330 protein.
GN At2G32330.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;

```

Score:	88.50	Matches:	28
Percent Similarity:	35.87%	Conservative:	5
Best Local Similarity:	30.43%	Mismatches:	32
Query Match:	8.07%	Indels:	27
DB:	10	Gaps:	3

  

US-09-914-191-1 (1-598) x Q93ZL5 (1-457)	
QY 71	GTTCGAACACGCGCAACAAGGTGTTCTGCTTAAACAGAGTAAGATACACACCCCCCATCC 130
Db 267	ValProLeuAnGlnThrAsnGluGluAlaArgThrVal-----SerProLeuProLys 284
QY 131	ATCCCTTCCTTCCTCCCTGTTCCCTCCCTCCCAACTTGAGTTGTGTCATTCGCACCCAGTCCTCTGG 190
Db 285	ValProCysPheProGlyProProProThrTrpTrpProTyRAlaTrpAsnGlyValSerTrp 304
QY 191	GTGTGAGGATGCTACGCCACCTAAGCGAAGG-----AGCCCTGGG 329
Db 305	ThrIleLeuProPheTyRProProAlaTyRTrpSerCysProGlyValSerProGly 324
QY 233	AGGTGGAGGGTTCATGGTTAAGCACACCGAAGCTGAAGCGCAAAAGGTCAGCTGTC 292
Db 325	AlaTrpAsnSerPhe-----ThrTrpMetProGlnProAsnSer 337
QY 293	TTCACTAGTAATCTCTGGATGTTCTTCCAGAAAGC 328
Db 330	-----ThrTrpMetProGlnProAsnSer 337

  

RESULT 11	
Q9FK95	
ID Q9FK95	PRELIMINARY; PRT; 515 AA.
AC Q9FK95;	
DT 01-MAR-2001 (TrEMBLrel. 16, Created)	
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)	
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)	
DE Similarity to H-protein promoter binding factor.	
OS Arabidopsis thaliana (Mouse-ear cress).	
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;	
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi	

```

OX NCBI_TAXID=3702;
[1]
RN SEQUENCE FROM N.A.
RP RC STRAIN=COLUMBIA;
RX MEDLINE=98403884; PubMed=9734815;
RA Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.
RT Sequence features of the regions of 1,367,185 bp covered by 19
RT physically assigned pl and TAC clones.";
RL DNA Res.. 5-203-216(1998).
DR ENBL; AB012243; BAB0898.1; .
DR InterPro; IPR003851; Znf_Dof.
DR Pfam; PF02701; zf-Dof.1.
SQ SEQUENCE 515 AA; 56528 MW; E2B5BF1957DC33FF CRC64;

Alignment Scores:
Pred. No.: 0.569 Length: 515
Score: 88.50 Matches: 28
Percent Similarity: 35.87% Conservative: 5
Best Local Similarity: 30.43% Mismatches: 32
Query Match: 8.07% Indels: 27
DB: 10 Gaps: 3

US-09-914-191-1 (1-598) x Q9FK95 (1-515)
Qy 71 GTTCCAAACCGCAACAAGGTGTTCTCTTAACACGAGTAGATACACCACCCCATTCC 130
Db 267 ValProLeuAsnGlnThrAsnGluAlaGlyThrVal-----SerProLeuProLys 284

Qy 131 ATCCCTTCCTTCCTGTTCCCTCCCAACTTGAGTTGTGTCAITCGCACCGAGTGCCTGG 190
Db 285 ValProCysPheProGlyVproProProThrTripproTyrAlaTrpAsnGlyValSerTrp 304

```

Qy 191 GTGCTAGGATGTACAGCCACCTAAGCAAGG-----AGCCTGGG 232  
Db 305 ThrileuProPheTyrProProAlaTyrTrpSerCysProGlyValSerProGly 324  
Qy 233 AGTGGAGGGCTTGCATGGTTAAGCACACCAAGAACTGAAGCGCAAAAGGTCAGCTGC 292  
Db 325 AlaTrpAsnSerPhe----- 329  
Qy 293 TTCATCTAGAATCTCTGGATGTTCTTCCAGAAAGC 328  
Db 330 -----ThrTrpMetProGlnProAsnSer 337  
RESULT 12  
ID O89277 PRELIMINARY; PRT; 390 AA.  
AC O89277;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Non-structural protein.  
OS Bombyx mori cytoplasmic polyhedrosis virus (BmCPV).  
OC Viruses; dsRNA viruses; Reoviridae; Cypovirus.  
OX NCBI\_TaxID=31594;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H;  
RX MEDLINE=98321164; PubMed=9657006;  
RA Hagiwara K., Tomita M., Kobayashi J., Miyajima S., Yoshimura T.;  
RT "Nucleotide sequences of Bombyx mori cytoplasmic polyhedrosis virus  
segment 8";  
RL Biochem. Biophys. Res. Commun. 247:549-553 (1998).  
DR EMBL; AB016437; BAA32043.1; -.  
SQ SEQUENCE 390 AA; 44151 MW; B172F1D7E790E2B9 CRC64;

Alignment Scores:  
Pred. No.: 1.36 Length: 390  
Score: 85.00 Matches: 27  
Percent Similarity: 50.00% Conservative: 24  
Best Local Similarity: 26.47% Mismatches: 43  
Query Match: 7.69% Indels: 8  
DB: 12 Gaps: 4

US-09-914-191-1 (1-598) x O89277 (1-390)  
Qy 577 GTACAAGCCCAAGGAGATTACAGGGAACCTTNCACGACACAGAAATGAGATTGTTT 518  
Db 136 IleguSerGluGlyGluMetAsnAspThrProGlnAspAspGluAlaGluLysLysAsp 155  
Qy 517 TCTGAATTCAACATCAACTATAATAAT-----GAGCTGCCGATGTATAGAAAGG 467  
Db 156 ValGluIleArgMetAsnTrpSerGluGluIleValGluLeuProLysHisGlnGlu 175  
Qy 466 ACTGTGTGTATATGGCAGAGGTGGATGAGTATGATGACAAAGAAATTAAGTGCACAA 407  
Db 176 AsnValLeuValLeuSerLysProSerMetIleSerGluGluLeuLeuMetProThr 194  
Qy 406 GAAATGAAGAAAAAGATGGCAGTGCACCGGACAGGACAAAGCCAGTGCCTTCAC 347  
Db 195 AspMetGlu-----ValAlaThrProLysValLeuGluProProThrProLeuPro 211  
Qy 346 TGCATATCATCGGGGATGTTTCTGGAAGGAACATCCAGAGATTCTAGATGAAGACAGC 287  
Db 212 SerProIleIle-----ValAlaValSerSerGluSerProGlnValLysGluIleGluArg 230  
Qy 286 TGACCC 281  
Db 231 LeuPro 232

RESULT 13  
Q921R1 PRELIMINARY; PRT; 495 AA.  
ID Q921R1  
AC Q921R1;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Ectonucleoside triphosphate diphosphohydrolase 2.  
GN ENTPD2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC011241; AAH11241.1; -.  
DR MGD; MGI:1096863; Entpd2.  
DR InterPro; IPR000407; GDA1\_CD39\_NTPase.  
DR Pfam; PF01150; GDA1\_CD39\_1.  
DR PROSITE; PS01238; GDA1\_CD39\_NTPASE; UNKNOWN\_1.  
KW Hydrolase.  
SQ SEQUENCE 495 AA; 54319 MW; A76468A0CBF86AAC CRC64;  
Alignment Scores:  
Pred. No.: 1.59 Length: 495  
Score: 84.50 Matches: 42  
Percent Similarity: 37.80% Conservative: 20  
Best Local Similarity: 25.61% Mismatches: 69  
Query Match: 7.71% Indels: 33  
DB: 11 Gaps: 9  
US-09-914-191-1 (1-598) x Q921R1 (1-495)  
Qy 87 CAAGGTGTTCTGCTTAAACAGAGTAACATACACACCCCATCCATCCCTTCCTCCCTG 146  
Db 250 GlnArgLeuLeuAlaSerAlaLeuGlnIleHisArgPheHisProCysTrpProLysGly 269  
Qy 147 TTCCCTCCCACTTGGTTGTCTATTCGCACCATGTCCTGGTGGTAGGATGCTAC 206  
Db 270 TyrSerThrGlnValLeuLeuArgGluValTyrGlnSerPro-----CysThr 285  
Qy 207 AGCCACCTAAGCAAGGAGCCCTGGGAGGTGGGAGGCTTGCAATGTTAAGCACACACAGA 266  
Db 286 MetGlyGlnArgProGlnThrPheAsnSerSerAlaThrValSerLeuSerGlyThrSer 305  
Qy 267 ACTGAAGCGCAAAAGGTCAGCTGTCTTCATCT-----AGAATCTCTGATGTTCTCTC 320  
Db 306 AsnAlaAlaLeuCysArgAspLeuValSerGlyLeuPheAsnIleSerCysProPhe 325  
Qy 321 CAGAAAGCATCCCCGATGATATCGCAGTGCAGGGCAGCTGGCTTCTCTGTCGGGTC 380  
Db 326 -----SerGlnCysSerPheAsnGly----- 332  
Qy 381 ACTGCCATCTTTTCTTCTCCATTTCTGTGGCAGCTTAATTTCTTTTGTGCATCACTCA 440  
Db 333 -----ValPheGlnProProValAla---GlyAsnPhelIleAlaPhe----- 345  
Qy 441 TCACCTTCTGCGCATATCAACAGTCCTTTCTTATACATCGGCAGCTCATTTATTATAG 500  
Db 346 SerAlaPheTyrTyr-----ThrValAspPheLeu-----LysThrValMetGly 360  
Qy 501 TTGATGTTGAATTCAGAAACAAATCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 550  
Db 361 LeuProValGlyThrLeuLysGlnLeuGlu-----AspAlaThrGluThrThrCysAsn 378  
Qy 561 CTCCTTGGGCT 572  
Db 379 GlnThrTrpAla 382  
RESULT 14  
Q9LP79 PRELIMINARY; PRT; 343 AA.  
ID Q9LP79  
AC Q9LP79;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
 DE TIN15.7.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,  
 RA Shinn P., Altafi H., Bei Q., Chin C., Chio J., Choi E., Conn L.,  
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
 RA Len C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
 RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
 RA Ecker J.R.;  
 RT "Genomic sequence for Arabidopsis thaliana BAC TIN15 from chromosome  
 I.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Ecker J.R.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
 RA Khan S., Kim C., Altafi H., Bei Q., Chin C., Chio J., Choi E.,  
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
 RA Lee J., Len C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,  
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
 RA Theologis A., Ecker J.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC020889; AAF79694.1; --  
 SQ SEQUENCE 343 AA; 38371 MW; CE0EAC1EA62DB64B CRC64;

Alignment Scores:  
 Pred. No.: 2.56 Length: 343  
 Score: 82.50 Matches: 45  
 Percent Similarity: 33.16% Conservative: 18  
 Best Local Similarity: 23.68% Mismatches: 62  
 Query Match: 7.53% Indels: 65  
 DB: 10 Gaps: 6

US-09-914-191-1 (1-598) x Q9LPT9 (1-343)

QY 85 AACAGGTGTTCTGCTTAAC-----AGAGTAAGATACACCCCATCCAT 132  
 DB 3 SerLysAlaLeuGlyPheLeuProArgLeuArgPheSerProArgLeu 22  
 QY 133 CCCTTCCTCCTGTTCCCTCCCACTTCAGTGTGTCATTCGCCAGTCCCTGGT 192  
 DB 23 LeuSerLeuProProSerProAlaSerSerThrPheAlaThrArgHisLysLeuAsp 42  
 QY 193 GGTAGGATGCTACAGCCACCTAAGGCAAGAGCCCTGGGAGGTGGAGGGCTTCATGG 252  
 DB 43 SerArgGlnThr-----LeuLeuTrp 49  
 QY 253 TTAGCACACAGAACCTGACGCGCAAGAGGTGACGTCTTCATCTAGATCTCTGGAT 312  
 DB 50 ---AsnLysProGlnLeuSerArgValArgValAlaCysSerSerGlnSerAspSer 68  
 QY 313 GTTCTTCAGAAAGCATCCCGCATGATGTCAGATGCAAGGCA---CTGGCTTTTTC 369  
 DB 69 ArgProGluLysLysGlnSerAspLysSerAsnTyrAlaArgAlaGluLeuPheArgGly 88  
 QY 370 TGTTCGGGTCA-----  
 DB 89 LysSerGlySerValSerPheAsnGlyLeuThrHisGlnLeuValGluGluSerLysLeu 108

QY 381 ----- 381  
 DB 109 ValSerAlaProPheGlnGluLysGlySerPheLeuTrpValLeuAlaProValVal 128  
 QY 382 -----CTGCCATCTTTTTCCTTCATTTCTGTGGCAGCTTAATTT 423  
 DB 129 LeuLeuSerSerLeuLeuLeuProGlnPhePheLeuSerGlyLeuLeuGluAlaThrPhe 148  
 QY 424 CTTTGTGATCATCTTCATCCACTTCTGCATATCAACA-----CAGTCCCTT 471  
 DB 149 LysAsnAspThrValAlaGlyArgSerGluLeuValThrSerPheCysPheGluThrVal 168  
 QY 472 TCTATACATCGCGACGCTCATTTATTAGT 501  
 DB 169 PheTyrAlaGlyLeuAlaLeuPheLeuSer 178  
 RESULT 15  
 Q9HCL7 PRELIMINARY; PRT; 1628 AA.  
 AC Q9HCL7;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE KIAA1555 protein (Fragment).  
 GN KIAA1555.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20450683; PubMed=10997877;  
 RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;  
 RT XVIII. The complete sequences of 100 new cDNA clones from brain which  
 code for large proteins in vitro.";  
 RL DNA Res. 7:273-281(2000).  
 DR EMBL; AB046775; BAB13381.1; --  
 DR HSSP; P15822; IBBO.  
 DR InterPro; IPR000822; Znf\_C2H2.  
 DR Pfam; PF00096; zf\_C2H2; 2.  
 DR SMART; SM00355; Znf\_C2H2; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 2.  
 KW DNA-binding; Metal-binding; Zinc-finger.  
 FT NON\_TER 1  
 SQ SEQUENCE 1628 AA; 174427 MW; 4E23A34360A3D939 CRC64;

Alignment Scores:  
 Pred. No.: 3.51 Length: 1628  
 Score: 82.00 Matches: 43  
 Percent Similarity: 31.02% Conservative: 24  
 Best Local Similarity: 19.91% Mismatches: 77  
 Query Match: 7.48% Indels: 72  
 DB: 4 Gaps: 9

US-09-914-191-1 (1-598) x Q9HCL7 (1-1628)

QY 38 TTAATAAAAGACACTTCTGCCAAAGAACCATGTTCCAAACACCGCAACAGGTGTTCT 97  
 DB 1204 LeuAlaArgLysHisSerLeuThrLysAsnAsp---SerSerProGln---ArgCysSer 1221  
 QY 98 GCTTAACAGAGTAAGATACACCCCATCCCTTCCTTC----- 142  
 DB 1222 ProAlaArgGluProGlnAlaSerAlaProSerProGlyLeuHisValAspProGly 1241  
 QY 143 -----CCTGTTCCCTCCCAACTGAGTTGTGTGTCATTCGCACCGAGTGTCC--- 187  
 DB 1242 ArgGlyMetGlyProLeuProCysGlySerProArgLeuGlnLeuSerProLeuThrLeu 1261  
 QY 187 ----- 187

```
Db 1262 CysProLeuGlyArgGluLeuAlaProArgAlaHisValLeuSerLysLeuGluGlyThr 1281
Qy 188 -----TGGGTGGTAGGGATGCTACAG 208
Db 1282 ThrAspProGlyLeuProArgTy-SerProThrArgArgTrpSerProGlyGlnAlaGlu 1301
Qy 209 ---CCACCTAAGGCAAGAGCCCTGGAGGTGGAGGGCTTGCATGGTTAAGCACACCAG 265
Db 1302 SerProProArgSerAlaProProGlyLysTrp---AlaLeuAlaGlyProGlySerPro 1320
Qy 266 AACTGAAGCGCAAAAGGTCAGCTCTTCATCTAGAACTCTCGATGTTCTTCCTCCAGAA 325
Db 1321 SerAlaGlyGluHisGlyProGlyLeuGlyLeuAlaProArgValLeuPheProAla 1340
Qy 326 AGCATCCCCGATGATATCGCAGTCGACAGGCGCACTGGCTTTGT----- 367
Db 1341 ProLeuProHisLysLeuLeuSerArgSerProGluThrCysAlaSerProTrpLysAla 1360
Qy 368 -----CTGGTCCGGGTAC----- 382
Db 1361 GluSerArgSerProSerCysSerProGlyProAlaHisProLeuSerSerArgProPhe 1380
Qy 383 -----TGCCATCTTTTTCCTTCATTCTGT-----TGGCAGCTTAATTTCTTTGT 430
Db 1381 SerAlaLeuHisAspPheHisGlyHisIleLeuAlaArgThrGluGluAsnIlePheSer 1400
Qy 431 CATCACTTCATCCACCTTCTGCCATATCAACACAGTCCTTCCTATA 478
Db 1401 HisLeuProLeuHisSerGlnHisLeuThrArgAlaProCysProLeu 1416
```

Search completed: July 9, 2003, 12:50:28  
Job time : 53.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 13:00:53 ; Search time 1245 Seconds  
(without alignments)  
13978.697 Million cell updates/sec

Title: US-09-914-191-1  
Perfect score: 598  
Sequence: 1 ttggaatagttcttctgttta.....gggttagtcacagattgttg 598

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_hgt:\*

3: gb\_in:\*

4: gb\_em:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_sts:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vi:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_or:\*

21: em\_ov:\*

22: em\_pat:\*

23: em\_ph:\*

24: em\_pl:\*

25: em\_ro:\*

26: em\_sts:\*

27: em\_un:\*

28: em\_vi:\*

29: em\_vt:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rtd:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	597	99.8	598	6	AX034339	AX034339 Sequence
2	597	99.8	1208	9	BC001523	BC001523 Homo sapi
3	597	99.8	1208	9	BC001852	BC001852 Homo sapi
4	593.8	99.3	1233	9	AK000553	AK000553 Homo sapi
5	591.2	98.9	2159	9	AK000953	AK000953 Homo sapi
6	590.6	98.8	1216	9	HS0801637	AL136669 Homo sapi
7	577.6	96.6	2263	9	AK021663	AK021663 Homo sapi
8	504	84.3	1780	9	AK023355	AK023355 Homo sapi
9	447	74.7	155344	9	AC026407	AC026407 Homo sapi
10	406.4	68.0	416	9	HUMZD93F03	AF086486 Homo sapi
11	370.8	62.0	374	11	G38490	G38490 SHGC-58349
12	306.6	51.3	567	11	G37344	G37344 SHGC-57583
13	177.4	29.7	207877	10	AL645948	AL645948 Mouse DNA
14	167	27.9	318930	2	AC073495	AC073495 Mus muscu
15	78.4	13.1	202051	2	AC128452	AC128452 Rattus no
16	50.8	8.5	37919	8	SPCC63	AL049522 S.pombe c
17	50.2	8.4	18333	2	AC019777	AC019777 Drosophil
18	50.2	8.4	184682	3	AC092238	AC092238 Drosophil
19	50.2	8.4	266308	3	AE003646	AE003646 Drosophil
20	50.2	8.4	303043	3	DROSADH05	AE003411 Drosophil
21	48.4	8.1	7218	6	I66494	I66494 Sequence 14
22	44.2	7.4	23463	2	AC020869	AC020869 Mus muscu
23	43.8	7.3	1719	8	SCYGR024C	Z72809 S.cerevisia
24	42.2	7.1	125020	9	AF429315	AF429315 Homo sapi
25	41.4	6.9	95746	2	AC097111	AC097111 Oryza sat
26	41.4	6.9	146261	2	AC097112	AC097112 Oryza sat
27	41.2	6.9	94075	2	AC131508	AC131508 Strongylo
28	41	6.9	165327	2	AC120699	AC120699 Rattus no
29	40.8	6.8	160759	2	AC117082	AC117082 Dictyoste
30	40.6	6.8	171115	2	AC102661	AC102661 Mus muscu
31	40	6.7	159693	10	AL646042	AL646042 Mouse DNA
32	40	6.7	224076	2	AC084019	AC084019 Mus muscu
33	39.4	6.6	117432	9	AC108704	AC108704 Homo sapi
34	39.4	6.6	151552	2	AC023437	AC023437 Homo sapi
35	39.4	6.6	227197	2	AC126179	AC126179 Homo sapi
36	39.2	6.6	43795	9	AC000022	AC000022 Genomic s
37	39.2	6.6	85472	9	AC025735	AC025735 Homo sapi
38	39.2	6.6	157706	2	AC073560	AC073560 Homo sapi
39	39.2	6.6	178542	2	AC096064	AC096064 Rattus no
40	39.2	6.6	186838	9	AC006338	AC006338 Homo sapi
41	39	6.5	12940	3	AE001386	AE001386 Plasmodiu
42	38.8	6.5	230487	2	AC116390	AC116390 Mus muscu
43	38.6	6.5	209973	2	AC073755	AC073755 Mus muscu
44	38.4	6.4	10212	3	DDI14911	AJ314911 Dictyoste
45	38.4	6.4	106144	2	AC116425	AC116425 Dictyoste

ALIGNMENTS

RESULT 1  
AX034339  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AX034339  
Sequence 1 from Patent WO0050637.  
AX034339  
AX034339.1 GI:10303095

human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 598)  
Godson,C.M., Brady,H.R. and Martin,F.M.  
Identification of genes having a role in the presentation of  
diabetic nephropathy





```

QY 481 TCAGCAGCTCATTATTATAGTTGATGTTGAATTCAGAAAACAAATCTCATTCTGTCTG 540
Db 707 TCAGCAGCTCATTATTATAGTTGATGTTGAATTCAGAAAACAAATCTCATTCTGTCTG 648
QY 541 CTGNAAGAGTTCCCTGTAATCTCCCTGGGCTGTACTGGTGTAGTCCAGATTGTG 598
Db 647 CTGAAAGAGTTCCCTGTAATCTCCCTGGGCTGTACTGGTGTAGTCCAGATTGTG 590

RESULT 3
BC001852/c
LOCUS BC001852 1208 bp mRNA linear PRI 12-JUL-2001
DEFINITION Homo sapiens, Similar to hypothetical protein FLJ20546, clone
MGC:4529 IMAGE:3010614, mRNA, complete cds.
ACCESSION BC001852
VERSION BC001852.1 GI:12804814
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1208)
AUTHORS Strausberg R.
TITLE Direct Submission
JOURNAL Submitted (29-JAN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
contact: amadan@systemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Kettman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 10 Row: g Column: 20.
Location/Qualifiers
1. .1208
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:4529 IMAGE:3010614"
/tissue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC 17"
/lab_host="DH10B-R"
/notes="Vector: pOT7"
383. .904
/codon_start=1
/product="Similar to hypothetical protein FLJ20546"
/protein_id="AAH01852.1"
/db_xref="GI:12804815"
/translation="MPTVASQFASVYVWRDYFEDQPLLYPPGDFRVVVPNOTL
KYLRSRQADCHNNLYNVFWALIOOSGLTPVQAGROGLTAAKNNELFSEFN
FNWKEPMPYRGTVLWQKVDVMTKEIKLPTMEGKMAVTRTRTKVPLHCDIIGA
YWKHEPEILDEDS"

BASE COUNT 346 a 258 c 315 g 289 t
ORIGIN
Query Match 99.8%; Score 597; DB 9; Length 1208;
Best Local Similarity 99.8%; Pred. No. 2.7e-160;
Matches 597; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTGGAATAGTCTTGCTTTTAAATAAGTACTGCGATTAAAAAAGACACTTCTGCCA 60
Db 1187 TTGGAATAGTCTTGCTTTTAAATAAGTACTGCGATTAAAAAAGACACTTCTGCCA 1128
QY 61 AAGGAACCATGTTCCAAACCGCAACAGAGTGTCTGCTTTAAACAGAGTAAGATACACC 120

```

```

Db 1127 AAGGAACCATGTTCCAAACCGCAACAGAGTGTCTGCTTTAAACAGAGTAAGATACACC 1068
QY 121 ACCCCCATCCATCCCTTCCCTTCCCTTCCCTCCCAACTTCAGTTGTGTCAATTCGCACC 180
Db 1067 ACCCCCATCCATCCCTTCCCTTCCCTTCCCTCCCAACTTCAGTTGTGTCAATTCGCACC 1008
QY 181 AGTGTCTCGGTGGTAGGATGCTACAGCCACTAAGGCAAGAGCCCTGGGAGGTGGGA 240
Db 1007 AGTGTCTCGGTGGTAGGATGCTACAGCCACTAAGGCAAGAGCCCTGGGAGGTGGGA 948
QY 241 GGGCTTGCATGTTAAGCACACAGCACTGAAGCGCAAAAGGTCAGCTGTCTTCATCTA 300
Db 947 GGGCTTGCATGTTAAGCACACAGCACTGAAGCGCAAAAGGTCAGCTGTCTTCATCTA 888
QY 301 GAATCTCTGGATGTTCTCTCCAGAAAGCATCCCGATGATATCGCAGTCAAGGGCAGCTG 360
Db 887 GAATCTCTGGATGTTCTCTCCAGAAAGCATCCCGATGATATCGCAGTCAAGGGCAGCTG 828
QY 361 GCTTTGCTCTGGTCCGGGTCACTGCCATCTTTTCTTCCATTTCTGTTGGCAGCTTAA 420
Db 827 GCTTTGCTCTGGTCCGGGTCACTGCCATCTTTTCTTCCATTTCTGTTGGCAGCTTAA 768
QY 421 TTTCCTTTGTCATCACTTCATCCACCTTCTGCCATATCAACACAGTCCCTTTCTTATACA 480
Db 767 TTTCCTTTGTCATCACTTCATCCACCTTCTGCCATATCAACACAGTCCCTTTCTTATACA 708
QY 481 TCGGCAGCTCATTATTATATAGTTGATGTTCAATTCAGAAAACAAATCTCATTCTGTCTG 540
Db 707 TCGGCAGCTCATTATTATATAGTTGATGTTCAATTCAGAAAACAAATCTCATTCTGTCTG 648
QY 541 CTGNAAGAGTTCCCTGTAATCTCCCTGGGCTGTACTGGTGTAGTCCAGATTGTG 598
Db 647 CTGNAAGAGTTCCCTGTAATCTCCCTGGGCTGTACTGGTGTAGTCCAGATTGTG 590

RESULT 4
AK000553/c
LOCUS AK000553 1233 bp mRNA linear PRI 22-FEB-2000
DEFINITION Homo sapiens CDNA FLJ20546 fis, clone KAT11504.
ACCESSION AK000553
VERSION AK000553.1 GI:7020725
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens signet-ring cell carcinoma cell_line:KATO III cDNA to
mRNA, clone lib:KAT clone:KAT11504.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
TITLE Unpublished
JOURNAL 2 (bases 1 to 1233)
REFERENCE Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
AUTHORS Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
TITLE Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
JOURNAL University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail:cdnal@ims.u-tokyo.ac.jp,
Tel:81-3-5449-5286, Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source
1. .1233
/organism="Homo sapiens"
/db_xref="taxon:9606"

```



QY 482 CGGAGCTGATTTATATAGTTGATGTTGAATTCAGAAAAAATAATCTATCTTGTCTGC 541  
 Db 1307 CGGCGGCTCATTTATATAGTTGATGTTGAATTCAGAAAAAATAATCTATCTTGTCTGC 1248  
 QY 542 TGAAGAGTTCCTGTAATCTCCCTCGGCTTGACTGTTGTTAGTCCAGATTGTTG 598  
 Db 1247 TGAAGAGTTCCTGTAATCTCCCTCGGCTTGACTGTTGTTAGTCCAGATTGTTG 1191

RESULT 6  
 HSM801637/c 1216 bp mRNA linear PRI 10-MAR-2001  
 LOCUS Homo sapiens mRNA; cDNA DKFZ564B1172 (from clone DKFZ564B1172);  
 DEFINITION complete cds.

ACCESSION AL136669  
 VERSION AL136669.1 GI:12052861

KEYWORDS human.  
 SOURCE

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1216)  
 AUTHORS Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,  
 Ansoerge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,  
 Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,  
 Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,  
 Wambutt R., Korn B., Klein M., and Poustka A.  
 TITLE Toward a Catalog of Human Genes and Proteins: Sequencing and  
 Analysis of 500 Novel Complete Protein Coding Human cDNAs

JOURNAL Genome Res. 11 (3), 422-435 (2001)  
 PUBMED 11230166

REFERENCE 2 (bases 1 to 1216)  
 AUTHORS Blum H., Bauersachs S., Mewes H.W., Gassenhuber J. and Wiemann S.

TITLE Direct Submission  
 JOURNAL Submitted (15-DEC-2000) MIPS, Am Klopferspitze 18a, D-82152

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by LMU (Ludwig Maximilians University,  
 Munich/Germany) within the cDNA sequencing consortium of the German  
 Genome Project.

This clone (DKFZ564B1172) is available at the RZPD in Berlin.  
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further  
 information about the clone and the sequencing project is available  
 at <http://www.mips.biochem.mpg.de/proj/cDNA/>.

FEATURES  
 Location/Qualifiers

1..1216  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="5"  
 /clone="DKFZ564B1172"  
 /tissue\_type="brain"  
 /clone\_lib="364 (synonym: hfb2). Vector pAMP1; host  
 XL-2blue; sites NotI + SalI"  
 /dev\_stage="fetal".  
 21..917  
 /gene="DKFZ564B1172"  
 21..917  
 /gene="DKFZ564B1172"  
 /note="FLJ20546 fis"  
 /codon\_start=1  
 /product="hypothetical protein"  
 /protein\_id="CA866604.1"  
 /db\_xref="GI:12052862"

gene  
 CDS  
 /translations="MWGACKVQVHDSLATISITRLRYLRIGATMAKSFYVDFRAD  
 QTLACHVVRVLDGRNFRFAEKHNFAPKENDSRALQMTKCAQTWMEELDIYAG  
 DSDEYFVFRKTNWFRFRASFKFWHVASQFSSYFWRDYFEDQFLYPPGFDGRV  
 VVPSNLTLDKXLSWQADCHINNLYNTVFWALIQOGLTFVQAGRLQGLTAAADKNE  
 ILFSEFNINWNLPMYRKCTVLVQKQVDEWVTKIKLPTMEGKKMAYTRTKFVP  
 LHCQIDIGAFKWFPELIDEDS"  
 1193

polyA\_site 342 a 261 c 319 g 294 t  
 BASE COUNT

## ORIGIN

Query Match 98.8%; Score 590.6; DB 9; Length 1216;  
 Best Local Similarity 99.2%; Pred. No. 1.9e-158;  
 Matches 593; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TTGGAATAGTTCCTGTTTATATAAATAGTACTGCGAATTAATAAAGACACTTCTGCGCA 60

Db 1200 TTTTGATAGTTCCTGCTTATATAAATAGTACTGCGAATTAATAAAGACACTTCTGCGCA 1141

QY 61 AAGGAACCATGTTCCAAACACCGCAACAGGTGTTCTGCTTTAAACAGAGTAACATACACC 120

Db 1140 AAGAAACCATGTTCCAAACACCGCAACAGGTGTTCTGCTTTAAACAGAGTAACATACACC 1081

QY 121 ACCCCCATCATCCCTTCCCTGTTCCCTGTTCCCTCCCAACTTGAGTTGTGTCAATTCGCACC 180

Db 1080 ACCCCCATCATCCCTTCCCTGTTCCCTGTTCCCTCCCAACTTGAGTTGTGTCAATTCGCACC 1021

QY 181 AGTGTCTCGGTGGTGGATGCTACAGCCACCTAAGGCAAGGAGCCCTGGGAGGTGGGA 240

Db 1020 AGTGTCTCGGTGGTGGATGCTACAGCCACCTAAGGCAAGGAGCCCTGGGAGGTGGGA 961

QY 241 GGGCTTGCATGGTTAAGCACACAGCAACTGAAGCGCAAAAGGTCAGCTGCTTTCATCTA 300

Db 960 GGGCTTGCATGGTTAAGCACACAGCAACTGAAGCGCAAAAGGTCAGCTGCTTTCATCTA 901

QY 301 GAATCTCTGGATGTTCCCTTCCAGAAAGCATCCCGCATGATATCGCAGTGCAGGGCACTG 360

Db 900 GAATCTCTGGATGTTCCCTTCCAGAAAGCATCCCGCATGATATCGCAGTGCAGGGCACTG 841

QY 361 GCTTTGCTCTGGTCCGGGTCAGTCCCATCTTTTTCCTCCCATCTTCTTGTGGCAGCTTAA 420

Db 840 GCTTTGCTCTGGTCCGGGTCAGTCCCATCTTTTTCCTCCCATCTTCTTGTGGCAGCTTAA 781

QY 421 TTTCTTTTCTCATCACTTCATCCACCTTCGCGCATATCAACAGTCCCTTTCCTATACA 480

Db 780 TTTCTTTTCTCATCACTTCATCCACCTTCGCGCATATCAACAGTCCCTTTCCTATACA 721

QY 481 TCGGCACTCATTTATATATAGTTGATGTTGAATTCAGAAACAAAAATCTCATTTGTCTG 540

Db 720 TCGGCACTCATTTATATATAGTTGATGTTGAATTCAGAAACAAAAATCTCATTTGTCTG 661

QY 541 CTGNAAGAGTTCCTGTAATCTCCCTTGGGCTTGACTGGTGTAGTCCAGATTGTTG 598

Db 660 CTGCAAGAGTTCCTGTAATCTCCCTTGGGCTTGACTGGTGTAGTCCAGATTGTTG 603

## RESULT 7

AK021663/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@ri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing: Research Association for Biotechnology; cDNA library  
construction, 5' - & 3'-end one pass sequencing and clone selection:  
Helix Research Institute (supported by Japan Key Technology Center  
etc.) and Department of Virology, Institute of Medical Science,  
University of Tokyo.

## FEATURES

source  
1. .2263  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HEMBA1003893"  
/tissue\_type="whole embryo, mainly head"  
/clone\_lib="HEMBA1"  
/dev\_stage="embryo, 10 weeks"  
/notes="cloning vector: pME183FL3"  
4. .900  
/notes="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB13870.1"  
/db\_xref="GI:10432894"  
DTLACHVVRVLDGRNFRFAEKHNFAPKNDRAQLMKCAOTVMELEDIVIAYG  
translation="MVGACKVVKVHDSLATISITRLRYLRIGATMAKSEYVVRDFEAD  
QSEYFVKRKTWPKBRASKEWTVASQFASVYVYWRDYFEDQPLLYPPGFGRV  
VYVPSNQLKYLSWQADCHNNLYNFWALIQOSGLTPVOAQRLOGTTLAADRKE  
ILPSEFNINYNPPMVRKGTVLWQVDEWMTKEIKLPTMEGKKMAVTRTRTRKVP  
LHCDIIGDAPFKEHPHILDEDS"  
BASE COUNT 624 a 486 c 580 g 573 t

Query Match 96.6%; Score 577.6; DB 9; Length 2263;  
Best Local Similarity 98.8%; Pred. No. 1e-154;  
Matches 592; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 2 TGGAAATAGTTCTTGCTTTATAAAATAGTACTCGATTAAAAAAAGCACTTCTGCCAA 61  
DB 1184 TGGAAATAGTTCTTGCTTTATAAAATAGTACTCGATTAAAAAAAGCACTTCTGCCAA 1125  
QY 62 AGGAACCATGTTCCAAACCGCAACAAAGGTGTCTGCTTTAAACAGAGTAAAGTACACCA 121  
DB 1124 AGGAACCATGTTCCAAACCGCAACAAAGGTGTCTGCTTTAAACAGAGTAAAGTACACCA 1065  
QY 122 CCCCCATCCATCCCTTCTTCCCTGTTCCCTCCAC--TTGAGTGTGCTATTCCGAC 179  
DB 1064 CCCCCATCCATCCCTTCTTCCCTGTTCCCTCCACATTTTGTGTGCTATTCCAC 1005  
QY 180 CAGTGTCTGGGTGAGGATGCTCAGGCACCTAAGCAAGGAGCCCTGGGAGGTGGG 239  
DB 1004 CAGTGTCTGGGTGAGGATGCTCAGGCACCTAAGCAAGGAGCCCTGGGAGGTGGG 945  
QY 240 AGGGCTTGATGTTAAGCACACAGAACTGAAGCGCAAAAGGTCAGTGTCTTCTATCT 299  
DB 944 AGGGCTTGATGTTAAGCACACAGAACTGAAGCGCAAAAGGTCAGTGTCTTCTATCT 885  
QY 300 AGAATCTCGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGGGCACT 359  
DB 884 AGAATCTCGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGGGCACT 825  
QY 360 GGCTTTGCTGTCGGGTGTCAGTGCATCTTTTCTTCCATTTCTGTTGCGAGCTTA 419  
DB 824 GGCTTTGCTGTCGGGTGTCAGTGCATCTTTTCTTCCATTTCTGTTGCGAGCTTA 765  
QY 420 ATTTCTTTTGTGATCACTTCCATCCACCTCTGCGATATCAACAGTCCCTTTCTTATAC 479  
DB 764 ATTTCTTTTGTGATCACTTCCATCCACCTCTGCGATATCAACAGTCCCTTTCTTATAC 705  
QY 480 ATCGGAGCTCATATATATAGTGTGAATTCAGAAACAAAATCTCATTTCTGTCT 539  
DB 704 ATCGGAGCTCATATATATAGTGTGAATTCAGAAACAAAATCTCATTTCTGTCT 645  
QY 540 GCTGNAAGAGTTCCTGTAAATCTCCCTTGGGCTGTGACTGGTGTAGTCCAGATTTGTG 598

DB 644 GCTGCAAGAGTTCCTGTATATCTCCCTGGGCTTGTAGTGTAGTCCAGATTGTG 586  
RESULT 8  
AK023355/c  
LOCUS Homo sapiens cDNA FLJ13293 fis, clone OVARC1001188. PRI 01-AUG-2002  
DEFINITION AK023355  
ACCESSION AK023355  
VERSION AK023355.1 GI:10435255  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens ovary, tumor tissue cDNA to mRNA, clone lib:OVARC1  
clone:OVARC1001188.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Isogai,T., Ota,T., Hayaashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,  
Nishikawa,T., Nagai,K., Sugano,S., Shiratori,A., Sudo,H.,  
Wagatsuma,M., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H.,  
Sugawara,M., Takahashi,M., Chiba,Y., Ishida,S., Murakawa,K.,  
Ono,Y., Takiguchi,S., Watanabe,S., Kimura,K., Murakami,K.,  
Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A.,  
Nakamura,Y., Nagahari,K., Masuho,Y., Ninomiya,K. and Iwayanagi,T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 1780)  
Isogai,T. and Otsuki,T.  
Direct Submission  
Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,  
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
(E-mail:genomics@ri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert  
sequencing: Research Association for Biotechnology; cDNA library  
construction, 5' - & 3'-end one pass sequencing and clone selection:  
Helix Research Institute (supported by Japan Key Technology Center  
etc.) and Department of Virology, Institute of Medical Science,  
University of Tokyo.  
FEATURES  
source  
1. .1780  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="OVARC1001188"  
/tissue\_type="ovary, tumor tissue"  
/clone\_lib="OVARC1"  
/note="cloning vector: pME183FL3"  
42. .563  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB14540.1"  
/db\_xref="GI:10435256"  
/translation="MTHVASQFASVYFVWRDYFEDQPLLYPPGFGRVYVPSNQL  
KDYLSWQADCHNNLYNFWALIQOSGLTEVQAGRLQGLTAAADKNILPSEFNIN  
YNELPMYRKGTVLWQVDEWMTKEIKLPTMEGKKMAVTRTRTRKPVLPVLRHCDIIGDA  
FWKEHPHILDEDS"  
BASE COUNT 513 a 368 c 435 g 464 t  
ORIGIN  
Query Match 84.3%; Score 504; DB 9; Length 1780;  
Best Local Similarity 93.0%; Pred. No. 1.4e-133;  
Matches 555; Conservative 0; Mismatches 1; Indels 41; Gaps 1;  
QY 2 TGAATAGTCTTGTCTTTATAAAATAGTACTCGATTAAAAAGCACTTCTGCCAA 61  
DB 804 TGAATAGTCTTGTCTTTATAAAATAGTACTCGATTAAAAAGCACTTCTGCCAA 745  
QY 62 AGGAACCATGTTCCAAACCGCAACAAAGGTGTCTGCTTTAAACAGAGTAAAGTACACCA 121  
DB 744 AGGAACCATGTTCCAAACCGCAACAAAGGTGTCTGCTTTAAACAGAGTAAAGTACACCA 685  
QY 122 CCCCCATCCATCCCTTCTTCCCTGTTCCCTCCCACTTGAGTGTGCTTCGACCA 181



Genome Sequencing Center  
Department of Genetics  
Washington University  
St. Louis MO 63108, USA  
http://genome.wustl.edu/gsc  
mailto:est@watson.wustl.edu

NOTICE: This sequence represents the full insert of this cDNA. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived. We have tried to obtain double-stranded, or double chemistry sequence across the entire clone, but potentially, there are areas in the sequence where this level of coverage was not achieved. Nevertheless, we are confident of the accuracy of this sequence as all regions of low quality, as defined by PHRAP (P. Green, in preparation), were visually inspected and edited accordingly. The consensus quality values for this sequence have been submitted separately.

The location of this clone is unknown.

FEATURES  
source  
Location/Qualifiers  
1..416  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:357053"  
/clone\_lib="Scars\_fetal\_heart\_NbHH19W"  
BASE COUNT 113 a 91 c 109 g 103 t  
ORIGIN

Query Match 68.0%; Score 406.4; DB 9; Length 416;  
Best Local Similarity 99.8%; Pred. No. 1.4e-105;  
Matches 407; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTGGAATAGTCTTCTTTATAAATAGTACTGCGATTAAATAAAGCACTTCTGCCA 60  
DB 408 TTGGAATAGTCTTCTTTATAAATAGTACTGCGATTAAATAAAGCACTTCTGCCA 349  
QY 61 AAGGAACCATGTTCCAAACCGCAACAGAGTGTCTGCTTAAACAGAGTAAGATACACC 120  
DB 348 AAGGAACCATGTTCCAAACCGCAACAGAGTGTCTGCTTAAACAGAGTAAGATACACC 289  
QY 121 ACCCCCATCATCCCTTCTTCCCTGTTCCTCCCACTCCCACTTGAGTTGTGTCATCGCACC 180  
DB 288 ACCCCCATCATCCCTTCTTCCCTGTTCCTCCCACTCCCACTTGAGTTGTGTCATCGCACC 229  
QY 181 AGTGTCTCGGTGGTAGGATGTACAGCCACCTAAGGCAAGGAGCCCTGGAGGTGGGA 240  
DB 228 AGTGTCTCGGTGGTAGGATGTACAGCCACCTAAGGCAAGGAGCCCTGGAGGTGGGA 169  
QY 241 GGGCTTCATGTTAAGCACACAGCACTGAAGCGCAAGAGGTCAGCTGCTTCATCTA 300  
DB 168 GGGCTTCATGTTAAGCACACAGCACTGAAGCGCAAGAGGTCAGCTGCTTCATCTA 109  
QY 301 GAATCTCTGATGTTCTCTCCAGAAAGCATCCCGGATGATATCGCAGTCAAGGGCACTG 360  
DB 108 GAATCTCTGATGTTCTCTCCAGAAAGCATCCCGGATGATATCGCAGTCAAGGGCACTG 49  
QY 361 GCTTTGCTCGGTCCGGGTCACTGCCATCTTTTTCCTTCCATTTCTG 408  
DB 48 GCTTTGCTCGGTCCGGGTCACTGCCATCTTTTTCCTTCCATTTCTG 1

RESULT 11  
G38490  
LOCUS  
DEFINITION SHGC-58349 Human Homo sapiens STS genomic, sequence tagged site.  
ACCESSION G38490  
VERSION G38490.1 GI:3249264  
KEYWORDS STS.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

# REFERENCE

1 (bases 1 to 374)  
Myers, R.M.  
Human STSs (1998)  
Unpublished (1998)  
COMMENT

Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157259689  
Email: myers@shgc.stanford.edu  
Primer A: CAACACCGCAACCAAGGTGT  
Primer B: GCTTAAACCATGCAAGCCCTC  
STS size: 184  
PCR Profile:

Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9700

Protocol:

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Amplifiraq Gold Polymerase: 0.07 units/uL

Total Vol: 5 uL

Buffer: MgCl2: 2.5 mM

KCl: 50 mM

Tris-HCl: 10 mM

pH: 8.3

Prepared with primer pairs derived from W93044 -- Unigene.

## FEATURES

Location/Qualifiers

1..374

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/map="5"

/clone\_lib="Human"

75..258

75..94

primer\_bind complement(239..258)

primer\_bind 99 a 97 c 87 g 91 t

BASE COUNT

ORIGIN

Query Match 62.0%; Score 370.8; DB 11; Length 374;

Best Local Similarity 99.5%; Pred. No. 2.3e-95;

Matches 372; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTGGAATAGTCTTCTTTATAAATAGTACTGCGATTAAATAAAGCACTTCTGCCA 60

DB 1 TTGGAATAGTCTTCTTTATAAATAGTACTGCGATTAAATAAAGCACTTCTGCCA 60

QY 61 AAGGAACCATGTTCCAAACCGCAACAGAGTGTCTGCTTAAACAGAGTAAGATACACC 120

DB 61 AAGGAACCATGTTCCAAACCGCAACAGAGTGTCTGCTTAAACAGAGTAAGATACACC 120

QY 121 ACCCCCATCATCCCTTCTTCCCTGTTCCTCCCACTCCCACTTGAGTTGTGTCATCGCACC 180

DB 121 ACCCCCATCATCCCTTCTTCCCTGTTCCTCCCACTCCCACTTGAGTTGTGTCATCGCACC 180

QY 181 AGTGTCTCGGTGGTAGGATGCTACAGCCACCTAAGGCAAGGAGCCCTGGAGGTGGGA 240

DB 181 AGTGTCTCGGTGGTAGGATGCTACAGCCACCTAAGGCAAGGAGCCCTGGAGGTGGGA 240

QY 241 GGGCTTCATGTTAAGCACACAGCACTGAAGCGCAAGAGGTCAGCTGCTTCATCTA 300

DB 241 GGGCTTCATGTTAAGCACACAGCACTGAAGCGCAAGAGGTCAGCTGCTTCATCTA 300

QY 301 GAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCCAGTGCAGGCGCACTG 360  
 Db 301 GAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCCAGTGCAGGCGCACTG 360  
 QY 361 GCTTTGCTCTGGTC 374  
 Db 361 GCTTTGCTCTGGTC 374

RESULT 12  
 G37344  
 LOCUS 567 bp DNA linear STS 31-MAR-1998  
 DEFINITION SHGC-57583 Human Homo sapiens STS genomic, sequence tagged site.  
 G37344  
 VERSION G37344.1 GI:2996995  
 KEYWORDS STS.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 567)  
 AUTHORS Myers R.M.  
 TITLE Human STSs (1997)  
 JOURNAL Human STSs (1997)  
 COMMENT Unpublished (1997)

Contact: Richard M. Myers  
 Stanford Human Genome Center (SHGC)  
 Stanford University School of Medicine  
 Department of Genetics, M-344, Stanford, CA 94305, USA  
 Tel.: 415/7259687  
 Fax: 415/7259689  
 Email: myers@shgc.stanford.edu  
 Primer A: CAACACCGCAACAAGGTGT  
 Primer B: TCCTTGCTTAGTGCTGT  
 STS size: 150  
 PCR Profile:  
 Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds  
 Annealing: 60 degrees C for 30 seconds  
 Polymerization: 72 degrees C for 23 seconds  
 PCR Cycles: 30  
 Thermal Cycler: Perkin Elmer 9700  
 Protocol:  
 Template: 25 ng  
 Primer: each 1 uM  
 dNTPs: each 200 uM  
 Amplitaq Gold Polymerase: 0.07 units/ul  
 Total Vol: 5 ul

Buffer: MgCl2: 2.5 mM  
 KCl: 50 mM  
 Tris-HCl: 10 mM  
 pH: 8.3

Prepared with primer pairs derived from W72675 -- Unigene.

FEATURES  
 source Location/Qualifiers  
 1..567  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /map="5"  
 /clone\_lib="Human"  
 79..228  
 79..98  
 primer\_bind  
 primer\_bind complement(209..228)  
 137 a 144 c 119 g 161 t 6 others

BASE COUNT 137 a 144 c 119 g 161 t 6 others  
 ORIGIN

Query Match 51.3%; Score 306.6; DB 11; Length 567;  
 Best Local Similarity 79.0%; Pred. No. 6.5e-77;  
 Matches 440; Conservative 0; Mismatches 6; Indels 111; Gaps 2;

QY 2 TGAATAGTCTTGTCTTTTATAAAATAGTACTGCGATATAAAAAAGACACTTCTGCCAA 61  
 Db 6 TGAATAGTCTTGTCTTTTATAAAATAGTACTGCGATATAAAAAAGACACTTCTGCCAA 65  
 QY 62 AGAACCATGTTCCAAACAGCGCAAGGTGTTCTGCTTAAACAGAGTAAGATACACCA 121  
 Db 66 AGAACCATGTTCCAAACAGCGCAAGGTGTTCTGCTTAAACAGAGTAAGATACACCA 125  
 QY 122 CCCCCATCCATCCCTTCCCTTCCCTCCCAACTCCCAACTGAGTTGTCATTCGCACCA 181  
 Db 126 CCCCCATCCATCCCTTCCCTTCCCTCCCAACTGAGTTGTCATTCGCACCA 185  
 QY 182 GTGTCTCTGGTGTAGGATGCTACAGCACTTAAGGCAAGAGCCCTGGAGGTGGAG 241  
 Db 186 GTGTCTCTGGTGTAGGATGCTACAGCACTTAAGGCAAGAGCCCTGGAGGTGGAG 245  
 QY 242 GCCTTCATGTTAAGCACACAGAACTGAAGCGCAAGGGTCACGTCTCTCCTAG 301  
 Db 246 GCCTTCGA----- 253  
 QY 302 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGAGTGCAGGCGCACTGG 361  
 Db 254 -----TGG 256  
 QY 362 CTTTGTCTGGTCCGGGTCACTGCCATCTTTTCTCCATTTCTGTGGCAGCTTAAT 421  
 Db 257 CTTTGTCTGGTCCGGGTCACTGCCATCTTTTCTCCATTTCTGTGGCAGCTTAAT 316  
 QY 422 TTCTTTTGTATCATCTTATCATCTTCACTTCTGCACTTCAACACAGTCCCTTCTATACAT 481  
 Db 317 TTCTTTTGTATCATCTTATCATCTTCACTTCTGCACTTCAACACAGTCCCTTCTATACAT 376  
 QY 482 CGGCAGCTCATTTATTATAGTTGATGTTGAATTCAGAAAAACAAATCTCATTCTTG--TCT 539  
 Db 377 CGGCAGCTCATTTATTATAGTTGATGTTGAATTCAGAAAAACAAATCTCATTCTTGCTG 436  
 QY 540 GCTGNAAGAGTTCCTCG 556  
 Db 437 GCTGCAAGAGTTCCTCG 453

RESULT 13  
 AL645948  
 LOCUS 207877 bp DNA linear ROD 16-FEB-2002  
 DEFINITION Mouse DNA sequence from clone Rp23-298M7 on chromosome 11, complete sequence.  
 AL645948  
 VERSION AL645948.10 GI:18855237  
 KEYWORDS HTG.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 207877)  
 DIRECT SUBMISSION  
 TITLE Direct Submission  
 JOURNAL Submitted (16-FEB-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
 COMMENT On Feb 21, 2002 this sequence version replaced gi:18491398. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following



abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-298M7 is from the RPI-23 Mouse PAC Library

constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6  
This sequence is the entire insert of clone RP23-298M7. The true left end of clone RP23-29H5 is at 164150 in this sequence. The right end of clone RP23-302J15 is at 69639 in this sequence.

## FEATURES

```

source          1..207877
                 /organism="Mus musculus"
                 /db_xref="taxon:10090"
                 /chromosome="11"
                 /clone="RP23-298M7"
                 /clone_lib="RPI-23"
                 91735..91776
                 /note="Sequence from uni-directional dGTP big dye
                 terminator reads only."
                 20490..204554
                 /note="Sequence from uni-directional primer reads and dGTP
                 big dye terminator reads only."
                 204823..204903
                 /note="Sequence from overlapping clone RP23-29H5
                 (AL663031). Assembly confirmed by restriction digest."
                 206726..206753
                 /note="Sequence from overlapping clone RP23-29H5
                 (AL663031). Assembly confirmed by restriction digest."
BASE COUNT      57557 a 45850 c 45787 g 58683 t
ORIGIN
Query Match      29.7%; Score 177.4; DB 10; Length 207877;
Best Local Similarity 72.7%; Pred. No. 1.1e-39;
Matches 271; Conservative 0; Mismatches 96; Indels 6; Gaps 3;
QY 84 AAACAGGTGTTCTGCTTAACAGAGTAAGATACACACCCCATCCATCCCTCTCTCC 143
Db 172211 AAAAAAGGTGTTCTGCTTAAGCAGAACAGATACACACCCCATATATCCC-TCCTTCA 172269
QY 144 CTGTTCCCTCCCAACTGATGTGTGTCATTGCCACAGTGCTCTGGGTGGTAGGATGC 203
Db 172270 TCTGTCTCCACCATCTCCAGTCATGTCCTCTTACAGTGTCTTGGACACGAGTGC 172329
QY 204 TACAGGCACCTAAGGCAAGAGCCCTGGAGGTGGAGGCTTGCATGTTAAGCACACC 263
Db 172330 TTGAACCCACTAGGGTTAGGG- - -GAGACTCGGGAGCCTTGTGTGATTAGCATGCT 172385
QY 264 AGAAGTGAAGCGCAAAAGGTCAGCTGTCTTCATCTAGATCTCTGGATGTTCTTCCAG 323
Db 172386 AGAAGTGAAGCGGTGA-GGTCAAGTCTCTTTCGCCAGAAATCTCTGGGTGTTCTTCCAG 172444
QY 324 AAAGCATCCCGATGATATCGCAGTGAAGGGCACTGGCTTGTCTGTGTCGGGTCACT 383
Db 172445 AAGCATCCCATATAGATACAGATTCAGGCCACACCCGGGTCTTGTTCAGGCACG 172504
QY 384 GCCATCTTTTCTTCCATTTCTGTGGCAGCTTAATTTCTTTTGTATCATCTCATCC 443
Db 172505 GCCTTCTTTTCCCTTCCATTTCTGTGGCAGCTTCTTGTGTGTTCTGACTTCTTCC 172564
QY 444 ACCTTCTGCCATA 456
Db 172565 ACCTATAGACAAA 172577

```

## RESULT 14

AC073495

LOCUS

DEFINITION

SEQUENCE, 64 unordered pieces.

AC073495

AC073495 318930 bp DNA linear HTG 15-OCT-2000

Mus musculus chromosome 11 clone RP23-345K4, WORKING DRAFT

AC073495

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AC073495.12 GI:10801940  
HTG; HTGS PHASE1; HTGS\_DRAFT.  
Mus musculus

## REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 318930)  
Metzker, M.D., Lewis, L.R., Hume, J., Edwards, C., Harris, C., Dedrich, D., Thomas, S., Okwuonu, G., Carlock, C., Garner, T., Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J., Buhatz, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z., Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S., Fernandez, C., Ferraguto, D., Forcum-Tansey, J., Gill, R., Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hognes, M., Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S., Kovar, C., Liu, J., Liu, W., Louisedge, H., Lozado, R.J., Martin, R., Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S., Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogih, M., Parish, B., Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S., Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A., Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstein, G., Worley, K. and Gibbs, R.  
Direct Submission

## TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information

Center project name: MACH

Center clone name: RP23-345K4

Summary Statistics

Sequencing vector: M13; L08821

Chemistry: Dye-terminator Big Dye; 71% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 282577 bases at least Q40

Consensus quality: 299603 bases at least Q30

Estimated insert size: 298164; sum-of-contigs estimation

Quality coverage: 0x in Q20 bases; agarose-fp estimation

Quality coverage: 5.1x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

NOTE: This sequence may represent more than one clone.

NOTE: This is a 'working draft' sequence. It currently

consists of 64 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

100449: contig of 100449 bp in length

100450: gap of unknown length

100550: contig of 43298 bp in length

143848: gap of unknown length

143948: contig of 8446 bp in length

152393: contig of 8446 bp in length

152394: gap of unknown length

152493: contig of 7933 bp in length

160426: gap of unknown length

160526: gap of unknown length

160527: contig of 6117 bp in length

166643: gap of unknown length

166743: gap of unknown length

173484: contig of 6741 bp in length



Query Match	Best Local Similarity	27.9%; Score 167; DB 2; Length 318930;
Matches 271; Conservative	72.7%; Pred. No. 1.1e-36;	Indels 7; Gaps 4;
QY	84	AAACAAGGTGTTCTGCTTAAACAGAGTAAGATACACACCCCATCCATCCCTCTCTCC 143
Db	258116	AAAAGAAGTGTTCCTTACGACAGACATACACACCCCATATATATCCC-TCCTTCA 258174
QY	144	CTGTTTCCCTCCCAACTTGAGTTGTGTCATTCGCAACAGTGTCTGGGTGGTAGGATGC 203
Db	258175	TCTGTCTCCACCATCTCCAGTCATGTCTCTTACAGTGTCTTGGGACAGGATGC 258234
QY	204	TACAGCCACTTAAGCAAGGCGCCCTGGGAGGTGGGAGGCTGCATGCTTAAGCACACC 263
Db	258235	TTGAACCACTTAGGGTTAGGGG----GAGACTCGGGAGCCTTGTCTGATTAAAGCATGT 258290
QY	264	AGAACTGAAGCGCAAAAGGGTCAGCTGTCTTTCATCTAGAACTCTCGATGTTTCTTCCAG 323
Db	258291	AGAACTGCAGAGCGTGA-GGTGAGTCTCTTCCGCCAGAACTCTGGGTGTTCTTCCAG 258349
QY	324	AAAGCATCCCCGATGATATCGCAGTGAAGGGCACTGGGCTTTGTCTGTGTCGGGTCACT 383

Db 258350 AAGGCATCCCAATAAGATCACAGTTCAAGGCCACACCGGGTCTGGTCCGAG-CACG 258408

QY 384 GCATCTTTTCTCCATTTCTGTTGGCAGCTTAATTTCTTTGTCATCATCTTCATCC 443  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 258409 GCCTTCTTTTCCCTTCCATTTCTGTTGGCAGCTTCATCTTGTGTCTGACTTCTTC 258468  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 444 ACCTTCTGCCATA 456

Db 258469 ACCTATAGACAAA 258481

RESULT 15  
AC128452

LOCUS AC128452.1 GI:21909148

DEFINITION Rattus norvegicus clone CH230-92F20, \*\*\* SEQUENCING IN PROGRESS

AC128452 202051 bp DNA linear HTG 19-JUL-2002

\*\*\*, 61 unordered pieces.

AC128452

VERSION AC128452.1

KEYWORDS HTG: HTGS PHASE1.

SOURCE Rattus norvegicus.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 202051)

AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaral, H.C., Are, J.R., Ayale, M., Banks, T., Barbara, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burrell, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrill, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Huly, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, D.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Monabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstein, G., and Gibbs, R.

TITLE Direct Submission

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 202051)

AUTHORS Worley, K.C.

TITLE Direct Submission

JOURNAL Submitted (19-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

## COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: GYWC  
Center clone name: CH230-92F20  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 151680 bases at least Q40  
Consensus quality: 160123 bases at least Q30  
Consensus quality: 167616 bases at least Q20  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 61 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1  
\* 1175: contig of 1175 bp in length  
\* 1176: gap of unknown length  
\* 1275: contig of 1074 bp in length  
\* 1276: gap of unknown length  
\* 2349: contig of 1074 bp in length  
\* 2350: gap of unknown length  
\* 2449: gap of unknown length  
\* 2450: contig of 1480 bp in length  
\* 3929: gap of unknown length  
\* 4029: gap of unknown length  
\* 4030: contig of 1240 bp in length  
\* 5269: gap of unknown length  
\* 5369: gap of unknown length  
\* 5370: contig of 1608 bp in length  
\* 6978: gap of unknown length  
\* 7077: gap of unknown length  
\* 7078: contig of 1643 bp in length  
\* 8721: gap of unknown length  
\* 8820: contig of 1681 bp in length  
\* 8821: gap of unknown length  
\* 10501: gap of unknown length  
\* 10502: contig of 1121 bp in length  
\* 11602: gap of unknown length  
\* 11722: gap of unknown length  
\* 11723: contig of 1278 bp in length  
\* 11822: gap of unknown length  
\* 11823: contig of 1278 bp in length  
\* 13101: gap of unknown length  
\* 13102: contig of 2029 bp in length  
\* 13201: gap of unknown length  
\* 13202: contig of 1105 bp in length  
\* 15330: gap of unknown length  
\* 15331: contig of 1105 bp in length  
\* 16434: gap of unknown length  
\* 16435: contig of 2234 bp in length  
\* 16535: gap of unknown length  
\* 18768: contig of 1011 bp in length  
\* 18769: gap of unknown length  
\* 18769: contig of 1011 bp in length  
\* 18869: gap of unknown length  
\* 18870: contig of 1722 bp in length  
\* 19880: gap of unknown length  
\* 19881: contig of 1722 bp in length  
\* 21701: gap of unknown length  
\* 21702: contig of 1858 bp in length  
\* 21802: gap of unknown length  
\* 23559: contig of 1858 bp in length  
\* 23560: gap of unknown length  
\* 23760: contig of 1976 bp in length  
\* 25735: gap of unknown length  
\* 25835: gap of unknown length  
\* 25836: contig of 1863 bp in length  
\* 27698: gap of unknown length  
\* 27699: contig of 1219 bp in length  
\* 27799: gap of unknown length  
\* 29017: contig of 1219 bp in length  
\* 29117: gap of unknown length  
\* 31521: contig of 2404 bp in length  
\* 31522: gap of unknown length  
\* 31621: gap of unknown length  
\* 33622: contig of 2041 bp in length  
\* 33623: gap of unknown length  
\* 33663: contig of 1822 bp in length  
\* 35584: gap of unknown length  
\* 35585: contig of 1827 bp in length  
\* 37511: gap of unknown length  
\* 37512: contig of 1798 bp in length  
\* 39409: gap of unknown length  
\* 39410: contig of 2730 bp in length  
\* 42239: gap of unknown length  
\* 39510

*	42240	42339: gap of unknown length
*	42340	44798: contig of 2459 bp in length
*	44799	44988: gap of unknown length
*	44899	47992: contig of 3094 bp in length
*	47993	48092: gap of unknown length
*	48093	51448: contig of 3356 bp in length
*	51449	51548: gap of unknown length
*	51549	54264: contig of 2716 bp in length
*	54265	54364: gap of unknown length
*	54365	57860: contig of 3496 bp in length
*	57861	57960: gap of unknown length
*	57961	60081: contig of 2121 bp in length
*	60082	60181: gap of unknown length
*	60182	63309: contig of 3128 bp in length
*	63310	63409: gap of unknown length
*	63410	66113: contig of 2704 bp in length
*	66114	66213: gap of unknown length
*	66214	69157: contig of 2944 bp in length
*	69158	69257: gap of unknown length
*	69258	72677: contig of 3420 bp in length
*	72678	72777: gap of unknown length
*	72778	76401: contig of 3624 bp in length
*	76402	76501: gap of unknown length
*	76502	79318: contig of 2817 bp in length
*	79319	79418: gap of unknown length
*	79419	82830: contig of 3412 bp in length
*	82831	82930: gap of unknown length
*	82931	84821: contig of 1891 bp in length
*	84822	84921: gap of unknown length
*	84922	89009: contig of 4088 bp in length
*	89010	89109: gap of unknown length
*	89110	92899: contig of 3790 bp in length
*	92900	92999: gap of unknown length
*	93000	97197: contig of 4198 bp in length
*	97198	97297: gap of unknown length
*	97298	100076: contig of 2779 bp in length
*	100177	100176: gap of unknown length
*	102841	102940: gap of unknown length
*	102941	108155: contig of 5215 bp in length
*	108156	108255: gap of unknown length
*	108256	112516: contig of 4261 bp in length
*	112517	112616: gap of unknown length
*	112617	115613: contig of 2996 bp in length
*	115613	115713: gap of unknown length
*	115713	118941: contig of 3229 bp in length
*	118942	119041: gap of unknown length
*	119042	122960: contig of 3919 bp in length
*	123061	123060: gap of unknown length
*	123061	126858: contig of 3798 bp in length
*	126859	126958: gap of unknown length
*	128959	132043: contig of 5084 bp in length
*	132043	132143: gap of unknown length
*	132143	137336: contig of 5494 bp in length
*	137337	137637: gap of unknown length
*	137637	142544: contig of 4808 bp in length
*	142545	142648: gap of unknown length
*	142648	148118: contig of 5474 bp in length
*	148119	148218: gap of unknown length
*	148219	152714: contig of 4496 bp in length
*	152715	152814: gap of unknown length
*	152815	158084: contig of 5270 bp in length
*	158085	158184: gap of unknown length
*	158185	162647: contig of 4463 bp in length
*	162648	162747: gap of unknown length
*	162748	170619: contig of 7872 bp in length
*	170620	170719: gap of unknown length

Db	80479	ACCTTTTGCACATCACTGTCCCTATCTCTGATCATATGTGGTTTCATGTGTGTAGTTG	80538
Qy	504	ATGTTGAATTCAGAAAAACAAATCTCATTTCTGTCTGTCGNAAGAGTTCCTCG	556
Db	80539	ATGTGAACTCGGAGAACAGATCTCATTTCTGTGTATGCTGTCTCAGAGTTTCCCTG	80591

Search completed: July 12, 2003, 14:05:19  
Job time : 1250 secs

Query Match 13.1%; Score 78.4; DB 2; Length 202051;  
Best Local Similarity 80.5%; Pred. No. 3e-11;  
Matches 91; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
Qy 444 ACCTTTGCGCATATCAACACAGTCCCTTTCCCTATACATCGCAGCTCATTTATATAGTTG 503

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 12:59:52 ; Search time 160 Seconds  
(without alignments)  
8416.844 Million cell updates/sec

Title: US-09-914-191-1  
Perfect score: 598  
Sequence: 1 ttggaatagttctgttcta.....gggttagtcagattgttg 598

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
N Geneseq_101002.*			
1:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*		
2:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*		
3:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*		
4:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*		
5:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*		
6:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*		
7:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*		
8:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*		
9:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*		
10:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*		
11:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*		
12:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*		
13:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*		
14:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*		
15:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*		
16:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*		
17:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*		
18:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*		
19:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*		
20:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*		
21:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*		
22:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*		
23:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*		
24:	/SID22/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	597	99.8	598	21	AA50408
2	597	99.8	1231	22	AA157967
3	594.4	99.4	1512	23	ABV2854
4	594.4	99.4	1512	23	ABV28683
5	591.2	98.9	2159	22	AAH13744
6	577.6	96.6	2263	22	AAH15315
7	541.6	90.6	557	22	AAH33902
8	504	84.3	1780	22	AAH16378
9	447	74.7	15865	22	AAK90825

C 10	364.4	60.9	618	22	AAH07151	Human cDNA clone (
C 11	344.4	57.6	1103	22	AA159753	Human polynucleoti
C 12	322.8	54.0	361	24	ABQ58609	Human colon cancer
C 13	259.8	43.4	390	23	ABV13617	Human prostate exp
C 14	237.4	39.7	856	22	AAH03726	Human cDNA clone (
C 15	202.6	33.9	287	23	ABV34731	Human prostate exp
C 16	202.6	33.9	287	23	ABV43586	Human prostate exp
C 17	164.2	27.5	290	23	ABV04448	Human cDNA clone (
C 18	125.6	21.0	563	22	AAH09037	Human cDNA clone (
C 19	117.2	19.6	545	24	ABQ58400	Human colon cancer
C 20	50.2	8.4	861	23	ABL07197	Drosophila melanog
C 21	50.2	8.4	2830	23	ABL06544	Drosophila melanog
C 22	50.2	8.4	2916	23	ABL07196	Drosophila melanog
C 23	39.2	6.6	43795	21	AAZ92583	Human DAZ genomic
C 24	39	6.5	5139	21	AAW70139	Plasmodium falcipa
C 25	38.6	6.5	1563	21	AAC45962	Arabidopsis thalia
C 26	37.2	6.2	7078	23	ABL03681	Drosophila melanog
C 27	37.2	6.2	13085	23	ABL03680	Drosophila melanog
C 28	37	6.2	580073	18	AAV58840	Mycoplasma genital
C 29	36.6	6.1	104	22	ABA72131	Human foetal liver
C 30	36.6	6.1	597	22	ABA59589	Human foetal liver
C 31	36.4	6.1	21729	23	ABL06154	Drosophila melanog
C 32	36.4	6.1	27423	23	ABL04268	Drosophila melanog
C 33	36.2	6.1	734	22	AAH36020	Human colon cancer
C 34	35.8	6.0	4084	12	AAQ13317	Duffy receptor gen
C 35	35.8	6.0	4084	16	AAQ83524	Duffy antigen bind
C 36	35.8	6.0	4084	18	AAT72889	DABP gene, Plasm
C 37	35.8	6.0	4084	21	AAZ98281	P. vivax DABP bind
C 38	35.8	6.0	5282	24	AA561406	Human gene regulat
C 39	35.6	6.0	451	19	AAV57516	Zcytor7 cytokine r
C 40	35.6	6.0	1963	22	ABA08826	Human class II cyt
C 41	35.6	6.0	1664976	19	AAV21209	Methanococcus jann
C 42	35	5.9	3343	23	ABL07990	Drosophila melanog
C 43	35	5.9	3782	23	ABL10494	Drosophila melanog
C 44	35	5.9	6660	22	AAH36164	Human musculoskele
C 45	35	5.9	6660	22	AAK71438	Human immune/haema

ALIGNMENTS

RESULT 1  
AA50408  
ID AAA50408 standard; cDNA; 598 BP.  
XX  
AC  
AA50408;  
XX  
DT 20-NOV-2000 (first entry)  
XX  
Human increased in high glucose IHG-1 cDNA.  
DE  
IHG-1; increased in high glucose 1; human; diabetic nephropathy;  
KW diabetes; differential expression; marker; diagnosis; ss.  
XX  
Homo sapiens.  
XX  
PN WO2000050637-A1.  
XX  
PD 31-AUG-2000.  
XX  
PF 28-FEB-2000; 2000WO-IE00026.  
XX  
PR 26-FEB-1999; 99IE-0000157.  
XX  
(HIBE-) HIBERGEN LTD.  
PA (UYDU-) UNIV COLLEGE DUBLIN.  
XX  
PI Brady HR, Godson CM, Martin FM;  
XX  
DR WPI; 2000-572102/53.  
XX  
PT Identifying genes used for the prevention and/or  
therapy of diabetic nephropathy involves culturing mesangial cells in

PT the presence of glucose which induces differential expression of  
 PT susceptible genes  
 XX

PS Claim 8; Page 7-8; 86pp; English.

XX The present sequence is that of a human IHG-1 (increased in high  
 CC glucose 1) cDNA partial sequence. IHG-1 was identified using a  
 CC novel method for identifying genes that have a role in the  
 CC presentation of diabetic nephropathy (DN). This method involves  
 CC culturing mesangial cells in a medium in the presence of a  
 CC concentration of glucose sufficient to induce differential  
 CC expression of a gene susceptible to such differential expression.  
 CC and identifying the gene so induced, especially by suppression  
 CC subtractive hybridisation. Further portions of IHG-1 cDNA are  
 CC given in AAM50409 and AAM50410. Genes identified by this method,  
 CC including IHG-1, can be used as a diagnostic marker for the  
 CC progression and presentation of DN, as an index of disease activity  
 CC and the rate of progression of DN, and as a basis for identifying  
 CC drugs for use in the prevention and/or therapy of DN (claimed).

XX SQ Sequence 598 BP; 145 A; 154 C; 122 G; 176 T; 1 other;

Query Match 99.8%; Score 597; DB 21; Length 598;

Best Local Similarity 100.0%; Pred. No. 4.4e-174;

Matches 598; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTGGAATAGTCTTCTTTTATAAATAGTACTGCGATTAAAAAAGACACTTCTGCCA 60  
 DB 1 TTGGAATAGTCTTCTTTTATAAATAGTACTGCGATTAAAAAAGACACTTCTGCCA 60  
 QY 61 AAGGAACCATGTTTCCAAACCGCAACAGGTGTTCTGCTTAAACAGAGTAAGATACACC 120  
 DB 61 AAGGAACCATGTTTCCAAACCGCAACAGGTGTTCTGCTTAAACAGAGTAAGATACACC 120  
 QY 121 ACCCCATCATCCCTTCTTCCCTGTCTCCCTCCCAACTTGAGTTGTGTCATTCGCACC 180  
 DB 121 ACCCCATCATCCCTTCTTCCCTGTCTCCCTCCCAACTTGAGTTGTGTCATTCGCACC 180  
 QY 181 AGTGTCTGGTGGTAGGATGCTACAGCCACCTAAGCCAGGAGCCCTGGAGGTGGGA 240  
 DB 181 AGTGTCTGGTGGTAGGATGCTACAGCCACCTAAGCCAGGAGCCCTGGAGGTGGGA 240  
 QY 241 GGGCTTGCATGTTAAGCACACACCAAGTGAAGCGCAAAAGGTCAGCTGTTCTATCTA 300  
 DB 241 GGGCTTGCATGTTAAGCACACCAAGTGAAGCGCAAAAGGTCAGCTGTTCTATCTA 300  
 QY 301 GAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGGGCACTG 360  
 DB 301 GAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGGGCACTG 360  
 QY 361 GCTTTGCTCTGGTCCGGTCACTGCCATCTTTTCTTCCATTTCTGTGGCAGCTTAA 420  
 DB 361 GCTTTGCTCTGGTCCGGTCACTGCCATCTTTTCTTCCATTTCTGTGGCAGCTTAA 420  
 QY 421 TTTCTTTTGTATCATCTTCAATCCACCTTCTGCGCATATCAACAGTCCCTTCTTATACA 480  
 DB 421 TTTCTTTTGTATCATCTTCAATCCACCTTCTGCGCATATCAACAGTCCCTTCTTATACA 480  
 QY 481 TCGGCAGCTCATTTATATAGTTGATGTTGAATTCAGAAACAAATCTCATTTCTGTCTG 540  
 DB 481 TCGGCAGCTCATTTATATAGTTGATGTTGAATTCAGAAACAAATCTCATTTCTGTCTG 540  
 QY 541 CTGNAAGAGTTCCTGTATCTCCCTGGCTGTGACTGGTGTAGTCCAGATTGTTG 598  
 DB 541 CTGNAAGAGTTCCTGTATCTCCCTGGCTGTGACTGGTGTAGTCCAGATTGTTG 598

RESULT 2

AA157967/c

ID AA157967 standard; cDNA; 1231 BP.

XX AC AA157967;

XX

DT

22-OCT-2001 (first entry)

XX

DE Human polynucleotide SEQ ID NO 170.

XX

Human; neurotrophic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KW peripheral nervous system; neuropathy; central nervous system; CNS;  
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 KW leukaemia; ss.

OS Homo sapiens.

XX WO200153312-A1.

PN

26-JUL-2001.

XX

26-DEC-2000; 2000WO-US34263.

XX

21-JAN-2000; 2000US-0488725.

PR

25-APR-2000; 2000US-0552317.

PR

09-JUL-2000; 2000US-0598042.

PR

19-JUL-2000; 2000US-0620312.

PR

03-AUG-2000; 2000US-0653450.

PR

14-SEP-2000; 2000US-0662191.

PR

19-OCT-2000; 2000US-0693036.

PR

29-NOV-2000; 2000US-0727344.

XX

(HYSE-) HYSEQ INC.

XX

Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

PI

Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI

Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX

WPI; 2001-442253/47.

DR

P-PSDB; AAM38811.

XX

Novel nucleic acids and polypeptides, useful for treating disorders

DR

such as central nervous system injuries -

PT

Claim 1; SEQ ID NO 170; 10078pp; English.

XX

The invention relates to human nucleic acids (AA157798-AA161369) and  
 the encoded polypeptides (AAM38642-AAM42213) with neurotrophic,  
 immunosuppressant and cytostatic activity. The polynucleotides are useful  
 in gene therapy. A composition containing a polypeptide or polynucleotide  
 of the invention may be used to treat diseases of the peripheral nervous  
 system, such as peripheral nervous injuries, peripheral neuropathy and  
 localized neuropathies and central nervous system diseases, such as  
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 utilisation of the activities such as: immune system suppression,  
 Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 assays for receptor activity, arthritis and inflammation, leukaemias and  
 C.N.S disorders.

CC

Note: The sequence data for this patent did not form part of the printed  
 specification.

CC

Sequence 1231 BP; 338 A; 273 C; 320 G; 300 T; 0 other;

XX

Query Match 99.8%; Score 597; DB 22; Length 1231;

Best Local Similarity 99.8%; Pred. No. 6.3e-174;

XX

Matches 597; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY

1 TTGGAATAGTCTTCTTTTATAAATAGTACTGCGATTAAAAAAGACACTTCTGCCA 60  
 |||||  
 1221 TTGGAATAGTCTTCTTTTATAAATAGTACTGCGATTAAAAAAGACACTTCTGCCA 1162

DB

61 AAGGAACCATGTTTCCAAACCGCAACAGGTGTTCTGCTTAAACAGAGTAAGATACACC 120  
 |||||  
 1161 AAGGAACCATGTTTCCAAACCGCAACAGGTGTTCTGCTTAAACAGAGTAAGATACACC 1102



KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 PN WO200160860-A2.  
 XX  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US05171.  
 XX  
 PR 17-FEB-2000; 2000US-183319P.  
 PR 16-MAR-2000; 2000US-189862P.  
 PR 25-MAY-2000; 2000US-207454P.  
 PR 03-JUN-2000; 2000US-211314P.  
 PR 18-JUL-2000; 2000US-219007P.  
 PR 13-DEC-2000; 2000US-255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 DR WPI; 2001-662795/76.  
 XX  
 XX Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer  
 XX  
 XX Claim 1; Page 6017; 11750pp; English.  
 XX  
 XX The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
 CC  
 XX Sequence 1512 BP; 407 A; 327 C; 401 G; 356 T; 21 other;  
 SQ  
 Query Match 99.4%; Score 594.4; DB 23; Length 1512;  
 Best Local Similarity 99.7%; Pred. No. 4.4e-173;  
 Matches 595; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 2 TGAATAGTCTTGGCTTTTATAAAATAGTACTCGATTAAAAAAGCACTTCTCCCAA 61  
 1486 TGAATAGTCTTGGCTTTTATAAAATAGTACTCGATTAAAAAAGCACTTCTCCCAA 1427  
 62 AGGAACCATGTTCCACACCGCAACAGGTGTTCTGCTTAACACAGTAAGATACACCA 121  
 1426 AGGAACCATGTTCCACACCGCAACAGGTGTTCTGCTTAACACAGTAAGATACACCA 1367  
 122 CCCCCATCCATCCCTTCCCTTCCCTCCCTCCCACTTGGTGTGTCATTGCGACCA 181  
 1366 CCCCCATCCATCCCTTCCCTTCCCTTCCCTCCCACTTGGTGTGTCATTGCGACCA 1307  
 182 GTGTCTCTGGTGTAGGATGCTACAGCCACCTTAAGCGAAGAGCCCTGGAGGTGGAG 241  
 1306 GTGTCTCTGGTGTAGGATGCTACAGCCACCTTAAGCGAAGAGCCCTGGAGGTGGAG 1247  
 242 GGCTTCATGGTTAAGCACACCAAGTGAAGCGCAAAAGGTCAGTCTTCATCTAG 301  
 1246 GGCTTCATGGTTAAGCACACCAAGTGAAGCGCAAAAGGTCAGTCTTCATCTAG 1187  
 302 AATCTCTGGATGTTCTTCCAGAAAGCATCCCCGATGATATCGAGTGAAGGGCACTGG 361

Db 1186 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGAGTGAAGGGCACTGG 1127  
 Oy 362 CTTTCTCTGGTCCGGTCCACATCTTTTCTCTCCATTTCTGTGGCAGCTTAAT 421  
 Db 1126 CTTTCTCTGGTCCGGTCCACATCTTTTCTCTCCATTTCTGTGGCAGCTTAAT 1067  
 Oy 422 TTCTTTTGTGCATCACTTTCATCCACCTTCTGCCATATCAACACAGTCCCTTTCTATACAT 481  
 Db 1066 TTCTTTTGTGCATCACTTTCATCCACCTTCTGCCATATCAACACAGTCCCTTTCTATACAT 1007  
 Oy 482 CGGCAGCTCATTTATTATAGTGTGAATTCAGAAAAAATACTTCTGTCTGC 541  
 Db 1006 CGGCAGCTCATTTATTATAGTGTGAATTCAGAAAAAATACTTCTGTCTGC 947  
 Oy 542 TGNAGAGTTCCTGTGAATCTCCCTGGCTTGTACTGGTGTAGTCAGATGTTG 598  
 Db 946 TCGAAGAGTTCCTGTGAATCTCCCTGGCTTGTACTGGCTTGTAGTCCAGATGTTG 890  
 RESULT 5  
 AAH13744/c  
 ID AAH13744 standard; cDNA; 2159 BP.  
 XX  
 AC AAH13744;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:10648.  
 XX  
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 KW Homo sapiens.  
 OS  
 XX EP1074617-A2.  
 PN  
 XX 07-FEB-2001.  
 PD  
 XX 28-JUL-2000; 2000EP-0116126.  
 PF  
 XX 29-JUL-1999; 95JP-0248036.  
 PR 27-AUG-1999; 95JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 PR  
 XX (HELI-) HELIX RES INST.  
 PA  
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 XX  
 XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 10648; 2537pp + CD ROM; English.  
 XX  
 XX The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and



CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX Sequence 2159 BP; 607 A; 463 C; 522 G; 567 T; 0 other;

Query Match 98.9%; Score 591.2; DB 22; Length 2159;  
 Best Local Similarity 99.3%; Pred. No. 5.1e-172;  
 Matches 593; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 TGGATAGTCTCTGTTTATTAATAATAGTACTGCGATTAAATAAAAGCACTTCTGCCAA 61  
 DB 1787 TGGATAGTCTCTGTTTATTAATAATAGTACTGCGATTAAATAAAAGCACTTCTGCCAA 1728

QY 62 AGGAACCATGTTTCCAAACCGCAACCAAGGTGTTCTGCTTAAACAGAGTAAGATACACCA 121  
 DB 1727 AGGAACCATGTTTCCAAACCGCAACCAAGGTGTTCTGCTTAAACAGATAGATACACCA 1668

QY 122 CCCCATCCATCCCTTCTCCCTGTTCCCTCCCACTGAGTTGTGTCATTCGCACCA 181  
 DB 1667 CCCCATCCATCCCTTCTCCCTGTTCCCTCCCACTGAGTTGTGTCATTCGCACCA 1608

QY 182 GTGTCTCGGTGGTAGGATGCTACAGCCACTAAGGCAAGAGCCCTGGGAGTGGAG 241  
 DB 1607 GTGTCTCGGTGGTAGGATGCTACAGCCACTAAGGCAAGAGCCCTGGGAGTGGAG 1548

QY 242 GGCTTGATGTTAAGCACACCAAGCACTGAAGCGCAAAAGGTCAGTGTCTTCATCTAG 301  
 DB 1547 GGCTTGATGTTAAGCACACCAAGCACTGAAGCGCAAAAGGTCAGTGTCTTCATCTAG 1488

QY 302 AATCTCGGATGTTCTTCCAGAAAGCATCCCGATGATATCGAGTGCAGAGGCACTGG 361  
 DB 1487 AATCTCGGATGTTCTTCCAGAAAGCATCCCGATGATATCGAGTGCAGAGGCACTGG 1428

QY 362 CTTTGTCTGTGTCGGGTCACTGCCATCTTTTTCCTTCCATTTCTGTGTCAGCTTAAT 421  
 DB 1427 CTTTGTCTGTGTCGGGTCACTGCCATCTTTTTCCTTCCATTTCTGTGTCAGCTTAAT 1368

QY 422 TTTCTTTGTGTCATCATCTTCATCCACCTTCTGTCATATCAACACAGTCCCTTTCCTATACAT 481  
 DB 1367 TTTCTTTGTGTCATCATCTTCATCCACCTTCTGTCATATCAACACAGTCCCTTTCCTATACAT 1308

QY 482 CGGAGCTCATTTATATAGTTGATGTTGAATTCAGAAAAAATAATCTCATTTGTCTGC 541  
 DB 1307 CGGAGCTCATTTATATAGTTGATGTTGAATTCAGAAAAAATAATCTCATTTGTCTGC 1248

QY 542 TGAAGAGTCCCTGTAATCTCCCTGGGCTTGACTGGTGTAGTCCAGATTGTTG 598  
 DB 1247 TGAAGAGTCCCTGTAATCTCCCTGGGCTTGACTGGTGTAGTCCAGATTGTTG 1191

## RESULT 6

AAH15315/c  
 ID AAH15315 standard; cDNA; 2263 BP.  
 XX  
 AC AAH15315;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA sequence SEQ ID NO:13474.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EPI074617-A2.  
 XX

PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Iehli S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 13474; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesising polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

Query Match 96.6%; Score 577.6; DB 22; Length 2263;  
 Best Local Similarity 98.8%; Pred. No. 8.4e-168;  
 Matches 592; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 2 TGAATAGTCTTGTCTTTTATAAAATAGTACTGCGATTAAATAAAAGCACTTCTGCCAA 61  
 DB 1184 TGAATAGTCTTGTCTTTTATAAAATAGTACTGCGATTAAATAAAAGCACTTCTGCCAA 1125

QY 62 AGGAACCATGTTCCAAACCGCAACCAAGGTGTTCTGCTTAAACAGAGTAAGATACACCA 121  
 DB 1124 AGGAACCATGTTCCAAACCGCAACCAAGGTGTTCTGCTTAAACAGAGTAAGATACACCA 1065

QY 122 CCCCATCCATCCCTTCTCCCTGTTCCCTCCCAAC--TTGAGTGTGTGTCATTTCGCAC 179  
 DB 1064 CCCCATCCATCCCTTCTCCCTGTTCCCTCCCAACCTTTTGTGTGTCATTTCACAC 1005

QY 180 CAGTGTCTGGTGGTAGGATGCTACAGCCACTTAAGCAAGAGCCCTGGGAGTGGG 239  
 DB 1004 CAGTGTCTGGTGGTAGGATGCTACAGCCACTTAAGCAAGAGCCCTGGGAGTGGG 945

QY 240 AGGCTTGCATGGTTAAGCACACCAAGCACTGAAGCCCAAAAGGTCAGTGTCTTCATCT 299  
 DB 944 AGGCTTGCATGGTTAAGCACACCAAGCACTGAAGCCCAAAAGGTCAGTGTCTTCATCT 885



PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX  
 XX WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs.  
 XX  
 PS Claim 8; SEQ ID 15323; 2537pp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 1780 BP; 513 A; 368 C; 435 G; 464 T; 0 other;

Query Match 84.3%; Score 504; DB 22; Length 1780;  
 Best Local Similarity 93.0%; Pred. No. 4.1e-145;  
 Matches 555; Conservative 0; Mismatches 1; Indels 41; Gaps 1;

QY 2 TGGAAATAGTTCTTCTTTATAAAATAGTACTGCGATTAAAAAGCACTTCTGCCAA 61  
 DB 804 TGGAAATAGTTCTTCTTTATAAAATAGTACTGCGATTAAAAAGCACTTCTGCCAA 745  
 QY 62 AGGAACCATGTTCCACACCGCAACAGGTTCTCTTAAACAGAGTAGATACACCA 121  
 DB 744 AGGAACCATGTTCCACACCGCAACAGGTTCTCTTAAACAGAGTAGATACACCA 685  
 QY 122 CCCCACATCCATCCCTTCTTCCCTGTTCCTCCCTCCCACTTGAGTTGTCTATTCGCACCA 181  
 DB 684 CCCCACATCCATCCCTTCTTCCCTGTTCCTCCCTCCCACTTGAGTTGTCTATTCGCACCA 625  
 QY 182 GTGTCCTGGTGGTAGGATGCTACAGCCACCTTAAGGCAAGGAGCCCTGGGAGTGGGAG 241  
 DB 624 GTGT-----CCTGGGAGTGGGAG 606  
 QY 242 GGCTTGATGGTTAAGCACACACCACTGAAGCCCAAAAGGTCAGTGTCTTCTATCTAG 301  
 DB 605 GGCTTGATGGTTAAGCACACCACTGAAGCCCAAAAGGTCAGTGTCTTCTATCTAG 546  
 QY 302 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGAAGGCACTGG 361  
 DB 545 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGAAGGCACTGG 486  
 QY 362 CTTTGTCTCGTCCGGGTCACTGCCATCTTTTCTTCTTCCATTTCTGTGTCAGCTTAAT 421

Db 485 CTTTGTCTCGTCCGGGTCACTGCCATCTTTTCTTCTTCCATTTCTGTGGCAGCTTAAT 426  
 QY 422 TTCTTTTGTTCATCATTCACTCCACCTTCTGCCATATCAACACAGTCCCTTTCTCTATACAT 481  
 Db 425 TTCTTTTGTTCATCATTCACTCCACCTTCTGCCATATCAACACAGTCCCTTTCTCTATACAT 366  
 QY 482 CGGCAGCTCAATTATATAGTTGATGTTGAATTCAGAAAAAATAATCTCTTGTCTGTCG 541  
 Db 365 CGGCAGCTCAATTATATAGTTGATGTTGAATTCAGAAAAAATAATCTCTTGTCTGTCG 306  
 QY 542 TGAAGAGTTCCCTCTGTAATCTCCCTTGGGCTTGTACTGGTGTAGTCCAGATTGTTG 598  
 Db 305 TGAAGAGTTCCCTCTGTAATCTCCCTTGGGCTTGTACTGGTGTAGTCCAGATTGTTG 249  
 RESULT 9  
 AAK90825/c  
 ID AAK90825 standard; DNA; 15865 BP.  
 XX  
 AC AAK90825;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human digestive system antigen genomic sequence SEQ ID NO: 4401.  
 XX  
 KW Human; digestive system antigen; Gene therapy; cancer; appendicitis;  
 KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
 KW digestive system disorder; Meckel's diverticulum; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155314-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01324.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232337.  
PR 14-SEP-2000; 2000US-0232338.  
PR 14-SEP-2000; 2000US-0232339.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.

PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
(HUMA-) HUMAN GENOME SCI. INC.  
PA  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-502630/55.  
XX  
Polynucleotides encoding digestive system antigens, useful for  
diagnosing, treating, preventing and/or prognosing disorders of the  
digestive system, particularly cancer and cancer metastases -  
Disclosure; SEQ ID NO 4401; 986pp; English.  
XX  
The present invention provides the protein and coding sequences of a  
number of human digestive system antigens. These can be used in the  
diagnosis, treatment and prevention of digestive system disorders,  
including cancer, Meckel's diverticulum, bacterial or parasitic  
infections, appendicitis, Hirschsprung's disease, chronic colitis or  
ulcerative colitis. The present sequence is a genomic DNA fragment  
encoding a digestive system antigen of the invention.  
XX  
Sequence 15865 BP; 4775 A; 3204 C; 3309 G; 4577 T; 0 other;  
SQ  
Query Match 74.7%; Score 447; DB 22; Length 15865;  
Best Local Similarity 98.9%; Pred. No. 5.1e-127;  
Matches 450; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 2 TGGATAGTCTTGTCTTTTATAAAATAGTACTGCGATTAAAAAAGACACTTCTGCCAA 61  
DB 14717 TGGATAGTCTTGTCTTTTATAAAATAGTACTGCGATTAAAAAAGACACTTCTGCCAA 14658  
QY 62 AGGAACCATGTTCCACACCGCAACCAAGGTGTTCTGCTTAAACAGAGTAAAGATACACCA 121  
DB 14657 AGGAACCATGTTCCACACCGCAACCAAGGTGTTCTGCTTAAACAGAGTAAAGATACACCA 14598  
QY 122 CCCCCATCCATCCCTTCCTTCCTTCCTCCCTCCCACTTGGTGTGTCATTCGCACCA 181  
DB 14597 CCCCCATCCATCCCTTCCTTCCTTCCTCCCTCCCACTTGGTGTGTCATTCGCACCA 14538  
QY 182 GTGTCTCTGGTGGTAGGATGCTACAGCCACCTTAAGCAAGAGCCCTGGGAGTGGGAG 241  
DB 14537 GTGTCTCTGGTGGTAGGATGCTACAGCCACCTTAAGCAAGAGCCCTGGGAGTGGGAG 14478  
QY 242 GCGTTCATGGTTAAGCACACAGAACTGAAGCGGAAAAGGTCACCTGTCTTCCTCTAG 301  
DB 14477 GCGTTCATGGTTAAGCACACAGAACTGAAGCGGAAAAGGTCACCTGTCTTCCTCTAG 14418

QY 302 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGGCGCACTGG 361  
 |||||  
 Db 14417 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGGCGCACTGG 14358  
 |||||  
 QY 362 CTTTGTCTGTCGGGTCACTGCATCTTTTTCCTTCCATTTCTGTGCGCATTAA 421  
 |||||  
 Db 14357 CTTTGTCTGTCGGGTCACTGCATCTTTTTCCTTCCATTTCTGTGCGCATTAA 14298  
 |||||  
 QY 422 TTTCTTTGTATCATCTTCATCCACCTTCTGCGATA 456  
 |||||  
 Db 14297 TTTCTTTGTATCATCTTCATCCACCTTCTGCGATA 14263  
 |||||

## RESULT 10

AAH07151/c

ID AAH07151 standard; cDNA; 618 BP.

XX AC AAH07151;

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA clone (5'-primer) SEQ ID NO:3986.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX XX 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

XX PR 02-MAY-2000; 2000JP-0183767.

XX PR 09-JUN-2000; 2000JP-0241899.

XX PA (HELI-) HELIX RES INST.

XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX PI Iehii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX DR WPI; 2001-318749/34.

XX PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 full-length cDNAs defined in the specification, and for the detection  
 and/or diagnosis of the abnormality of the proteins encoded by the  
 full-length cDNAs.

XX PS Claim 1; SEQ ID 3986; 2537pp + CD ROM; English.

XX CC The present invention describes primer sets for synthesising 5602  
 full-length cDNAs defined in the specification. Where a primer set  
 comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 to the complementary strand of a polynucleotide which comprises one of  
 the 5602 nucleotide sequences defined in the specification, where the  
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 of an oligonucleotide comprising a sequence complementary to the  
 complementary strand of a polynucleotide which comprises a 5'-end  
 sequence and an oligonucleotide comprising a sequence complementary to a  
 polynucleotide which comprises a 3'-end sequence, where the  
 oligonucleotide comprises at least 15 nucleotides and the combination of  
 the 5'-end sequence/3'-end sequence is selected from those defined in  
 the specification. The primer sets can be used in antisense therapy and  
 in gene therapy. The primers are useful for synthesising polynucleotides,  
 particularly full-length cDNAs. The primers are also useful for the  
 detection and/or diagnosis of the abnormality of the proteins encoded by  
 the full-length cDNAs. The primers allow obtaining of the full-length  
 cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.

XX SQ Sequence 618 BP; 186 A; 135 C; 145 G; 149 T; 3 other;

Query Match 60.9%; Score 364.4; DB 22; Length 618;

Best Local Similarity 98.6%; Pred. No. 3.4e-102;

Matches 365; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 229 TGGAGGTGGGGGCTTGCATGTTAAGCACACCAAGCACTGAAGCGCAAAAGGTCAGC 288

|||

Db 618 TGNAGGTGAAGGGCTTGCATGTTAAGCACACCAAGCACTGAAGCGCAAAAGGTCAGC 559

|||

QY 289 TGTCTTCATCTAGAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGT 348

|||

Db 558 TGTCTTCATCTAGAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGT 499

|||

QY 349 GCAAGGGCACTGGCTTTGCTGGTCCGGGTCACTGCCATCTTTTCTTCCATTTCGT 408

|||

Db 498 GCAAGGGCACTGGCTTTGCTGGTCCGGGTCACTGCCATCTTTTCTTCCATTTCGT 439

|||

QY 409 TTGCGAGCTTAATTTCTTTTGTTCATCATCTTCACCTTCGCCATATCAACACAGTCC 468

|||

Db 438 TTGCGAGCTTAATTTCTTTTGTTCATCATCTTCACCTTCGCCATATCAACACAGTCC 379

|||

QY 469 CTTTCTATACATCGGCAGCTCATTTATTATAGTTGATTTGAATTCAGAAAAAATAATCT 528

|||

Db 378 CTTTCTATACATCGGCAGCTCATTTATTATAGTTGATTTGAATTCAGAAAAAATAATCT 319

|||

QY 529 CATTCTGTCTGCTGNAAGAGTTCCCTGTAATCTCCCTTGGGCTTGTACTGGTGTAGTC 588

|||

Db 318 CATTCTGTCTGCTGCAAGAGTTCCCTGTAATCTCCCTTGGGCTTGTACTGGTGTAGTC 259

|||

QY 589 CAGATTGTTG 598

|||

Db 258 CAGATTGTTG 249

|||

RESULT 11

AAI59753/c

ID AAI59753 standard; cDNA; 1103 BP.

XX AC AAI59753;

XX DT 22-OCT-2001 (first entry)

XX DE Human polynucleotide SEQ ID NO 3742.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX KW peripheral nervous system; neuropathy; central nervous system; CNS;

XX KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

XX KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

XX KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

XX KW leukaemia; ss.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0488725.

XX PR 25-APR-2000; 2000US-0552317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX XX

PA (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 XX WPI: 2001-442253/47.  
 DR P-PSDB; AAM40597.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 XX Claim 1; SEQ ID NO 3742; 10078pp; English.  
 XX The invention relates to human nucleic acids (AAI57798-AAI51369) and  
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 XX Sequence 1103 BP; 303 A; 242 C; 290 G; 268 T; 0 other;  
 . SQ  
 Query Match 57.6%; Score 344.4; DB 22; Length 1103;  
 Best Local Similarity 81.1%; Pred. No. 6.9e-96;  
 Matches 486; Conservative 0; Mismatches 2; Indels 111; Gaps 3;  
 .  
 QY 2 TGGATATAGTCTTCTTATATAAATAGTACTGCGATTAAATAAAGCACTTCTGCCAA 61  
 Db 1098 TGGATATAGTCTTCTTATATAAATAGTACTGCGATTAAATAAAGCACTTCTGCCAA 1039  
 .  
 QY 62 AGGAACCATGTTTCAACACCGCAAAAGGTTTCTGCTTAAACAGAGTAAAGATACACCA 121  
 Db 1038 AGGAACCATGTTTCAACACCGCAAAAGGTTTCTGCTTAAACAGAGTAAAGATACACCA 979  
 .  
 QY 122 CCCCACATCATCCCTTCCCTTCCCTGTTCCCT-CCCAACTGTAGTTGTGTCATCGCACC 180  
 Db 978 CCCCACATCATCCCTTCCCTTCCCTGTTCCCTCCTCCCAACTGTAGTTGTGTCATCGCACC 919  
 .  
 QY 181 AGTGCTCTGGTGTAGGGATGCTACAGCCACCTAAGGCAAGGCC-TCGGAGGTGGG 239  
 Db 918 AGTGCTCTGGTGTAGGGATGCTACAGCCACCTAAGGCAAGGCC-TCGGAGGTGGG 859  
 .  
 QY 240 AGGCTTGCATGGTAAAGCACACAGAACTGAAGCGCAAAAGGTCAGCTGCTTCTATCT 299  
 Db 858 AGGCTTGCATGGTAAAGCACACAGAACTGAAGCGCAAAAGGTCAGCTGCTTCTATCT 849  
 .  
 QY 300 AGAATCTGTGATGTTCTTCCAGAAAGCATFCCCGATGATATCGCAGTCAAGGGCACT 359  
 Db 848 AGAATCTGTGATGTTCTTCCAGAAAGCATFCCCGATGATATCGCAGTCAAGGGCACT 848  
 .  
 QY 360 GGCTTTGCTGTGTCGGGTCACTGCCATCTTTTCTTCCATTTCTGTTGGCAGCTTA 419  
 Db 847 GGCTTTGCTGTGTCGGGTCACTGCCATCTTTTCTTCCATTTCTGTTGGCAGCTTA 788  
 .  
 QY 420 ATTTCTTTTGTGTCATCACTTTCATCACCCTTCTGCCATATCAACAGATCCCTTTCCCTATAC 479  
 Db 787 ATTTCTTTTGTGTCATCACTTTCATCACCCTTCTGCCATATCAACAGATCCCTTTCCCTATAC 728  
 .  
 QY 480 ATCGGAGCTCATTTATATAGTTGATGTTGAATTCAGAAAAACAAATCTCATTTCTGTCT 539  
 Db 727 ATCGGAGCTCATTTATATAGTTGATGTTGAATTCAGAAAAACAAATCTCATTTCTGTCT 668  
 .  
 QY 540 GCTGNAAGAGTTCCTCTGTAATCTCCCTGGGCTTGTACTGGTCTAGTCCAGATTGTTG 598  
 Db 667 GCTGCAAGAGTTCCTCTGTAATCTCCCTGGGCTTGTACTGGTCTAGTCCAGATTGTTG 609  
 .  
 RESULT 12  
 ABQ58609  
 ID ABQ58609 standard; cDNA; 361 BP.  
 XX AC ABQ58609;  
 XX DT 02-AUG-2002 (first entry)  
 XX DE Human colon cancer related nucleotide sequence SEQ ID NO:2304.  
 XX KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;  
 XX genetic analysis; diagnostic; antisense therapy; gene; ss.  
 XX OS Homo sapiens.  
 XX PN WO200229086-A2.  
 XX PD 11-APR-2002.  
 XX PF 02-OCT-2001; 2001WO-US30732.  
 XX PR 02-OCT-2000; 2000US-237271P.  
 XX PA (FARB ) BAYER CORP.  
 XX PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;  
 PI Thiagalingam A, Lewis ME;  
 XX WPI: 2002-426115/45.  
 XX New isolated nucleic acid that is differentially expressed in cancer  
 PT tissues useful for determining the presence of colon cancer in a cell  
 PT or tissue type, and in antisense therapy -  
 XX Claim 1; Fig 1; 796pp; English.  
 .  
 CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially  
 CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins  
 CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be  
 CC used in antisense therapy. An antibody immunoreactive with a polypeptide  
 CC encoded by (I) is useful for detecting cancer in a patient sample, and  
 CC for detecting the presence or absence of a polynucleotide encoded by a  
 CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived  
 CC from (I) can be used for determining the presence of a nucleic acid which  
 CC hybridises to (I), and for determining the phenotype of cells in a sample  
 CC of cells from a patient. (I) is useful for determining the presence of  
 CC colon cancer in a cell or tissue type, for determining the presence of  
 CC state of other type of cancer, in antisense therapy, to generate  
 CC macroarrays on a solid surface, to identify a chromosome on which the  
 CC corresponding gene resides, and in tissue profiling, forensics, genetic  
 CC analysis, mapping and diagnostic applications. (I) can be used to raise  
 CC antibodies, and to screen for peptide analogues and antagonists.  
 .  
 XX Sequence 361 BP; 91 A; 102 C; 81 G; 82 T; 5 other;  
 . SQ  
 Query Match 54.0%; Score 322.8; DB 24; Length 361;  
 Best Local Similarity 97.5%; Pred. No. 1.9e-89;  
 Matches 346; Conservative 0; Mismatches 7; Indels 2; Gaps 2;  
 .  
 QY 31 ACTGCGATTAAAAAAGCACTTCTGCCAAAGAACCATGTTTCCAAACCGCAACAAG 90  
 Db 1 ACTGCGATTAAAAAAGCACTTCTGCCAAAGAACCATGTTTCCAAACCGCAACAAG 60  
 .  
 QY 91 GTGTTCTGCTTAAACAGAGTAAAGATACACACCCCATCCCTTCTTCTTCTTCTTCC 150  
 Db 61 GTGTTCTGCTTAAACAGAGTAAAGATACACACCCCATCCCTTCTTCTTCTTCTTCC 120  
 .  
 QY 151 CTTCCCAACTTGAGTTGTGTCTTCATTCGACAGTGCTCTGGTGTAGGATGCTACAGCC 210

Db 121 CCTCCCACTTGGTGTGATTCGACCAAGTGTCTGGGTGTAGGATGCTACAGCC 180  
Qy 211 ACCTAAGCAAGAGAGCCCTGGAGGTGGAGGGCTTGATGTTAAGCACACCAAGAACTG 270  
Db 181 ACCTAAGCAAGAGAGCCCTGGAGGTGGAGGGCTTGATGTTAAGCACACCAAGAACTG 240  
Qy 271 AAGCGCAAAAGGGTCAGC-TGCTTTCATCTAGAAATCTCTGATGTTCTTCCAGAAAGCA 329  
Db 241 AAGCGCAAAAGGGTCAGCTTCTTTCATCTAGAAATCTCTGATGTTCTTCCAGAAAGCA 300  
Qy 330 TCCCGGATGATATCGCAGTG-CAAGGCACTGGCTTTGTCTGTCGGGTCACT 383  
Db 301 TCCCGGATGATATCGCAGTG-CAAGGCACTGGCTTTGTCTGTCGGGTCACT 355

## RESULT 13

ABV13617/c  
ID ABV13617 standard; cDNA; 390 BP.

XX AC ABV13617;

XX 13-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 13608.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US05171.

XX 17-FEB-2000; 2000US-183319P.

XX 16-MAR-2000; 2000US-189862P.

XX 25-MAY-2000; 2000US-207454P.

XX 09-JUN-2000; 2000US-211314P.

XX 18-JUL-2000; 2000US-219007P.

XX 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;  
XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer -

XX Claim 1; Page 2258; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

XX specification or its complement. (I) is useful for:

XX (a) assessing whether a patient is afflicted with prostate cancer;

XX (b) monitoring the progression of prostate cancer in a patient;

XX (c) assessing the efficacy of a test compound to inhibit prostate

XX cancer in a patient;

XX (d) assessing the efficacy of a therapy for inhibiting prostate cancer

XX in a patient;

XX (e) selecting a composition for inhibiting prostate cancer in a patient;

XX (f) assessing the prostate cell carcinogenic potential of a compound;

XX (g) determining whether prostate cancer has metastasized in a patient;

XX (h) assessing the aggressiveness or indolence of prostate cancer in a

XX patient;

XX (i) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 390 BP; 126 A; 74 C; 104 G; 86 T; 0 other;

XX SQ

Query Match 43.4%; Score 259.8; DB 23; Length 390;  
Best Local Similarity 87.9%; Pred. No. 5.7e-70;  
Matches 341; Conservative 0; Mismatches 38; Indels 9; Gaps 5;  
Qy 199 GATGCTACAGCACTTAAGGCAAGGAGCCCTGGGA--GGTGGAGGGCTTGCAATGGTTAA 256  
Db 390 GATGCTTCAACCCCTTAAGGCCAGAGCCCTTGAAGGTTGAAGGGCTTGCAATGGTTAA 331  
Qy 257 GCACA-CCGAACTGAAGCGCAAAAGGGTCAGCTG---TCTTCATCTAGAAATCTCTGGA 311  
Db 330 GCCCCCCCAGAACTTAAGCGCAAAAGGGTCAACTGTTTTTCATCTTAAATTTCTGGAAT 271  
Qy 312 TGTTCCTTCAGAAAGCATCCCCGATGATATCGCAGTGCAGAGGGCACTTGGCTTTGCTCTG 371  
Db 270 GTTCTCTTCCAAAAGCAATCCCAATGATATCGCAGTGCAGAGGGCCCTGGCTTTGCTCTG 211  
Qy 372 GTCCGGGTCACTGCCATCTTTTTTCTTCATTTCTGTTGGGAGCTTAA-TTCTTTTGT 430  
Db 210 GTCCGGGTCACTGCCATCTTTTTTCTTCATTTCTGTTGGGAGCTTAA-TTCTTTTGT 151  
Qy 431 CATCACTTCAT-CCACCTTCTGCATATCAACACAGTCCCTTTTCTTATACATCGGAGCT 489  
Db 150 TATCACTTCATCCCACTTCTGCCATATCAACACAGTCCCTTTTCTTATACATCGGAGCT 91  
Qy 490 CATTATTATAGTTGATGTTGAATTCAGAAAAACAAAATCTCATTTCTGTCTGCTGNAAGAG 549  
Db 90 CATTATTATAGTTGATGTTGAATTCAGAAAAACAAAATCTCATTTCTGTCTGCTGCAAGAG 31  
Qy 550 TTCCCTGTAAATCTCCCTTGGGCTTGATAC 577  
Db 30 TTCCCTGTAAATCTCCCTTGGGCTTGATAC 3

## RESULT 14

AAH03726/c  
ID AAH03726 standard; cDNA; 856 BP.

XX AC AAH03726;

XX 26-JUN-2001 (first entry)

XX Human cDNA clone (5'-primer) SEQ ID NO:561.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
XX full-length cDNAs defined in the specification, and for the detection  
XX and/or diagnosis of the abnormality of the proteins encoded by the  
XX full-length cDNAs -

XX Claim 1; SEQ ID 561; 2537pp + CD ROM; English.

XX PS



XX The present invention describes primer sets for synthesising 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB98993 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 856 BP; 248 A; 183 C; 222 G; 196 T; 7 other;

Query Match 39.7%; Score 237.4; DB 22; Length 856;  
Best Local Similarity 94.1%; Pred. No. 7.1e-63;  
Matches 255; Conservative 0; Mismatches 15; Indels 1; Gaps 1;  
QY 329 ATCCCCGATATCGCAGTGCAAGGCACTGGCTTGTCTCGTCCGGGTCACTGCCAT 388  
Db |||||||  
QY 856 ATCCCCGATATCGCAGTGCAAGGCACTGGCTTGTCTCGTCCGGGTCACTGCCAT 797  
Db |||||||  
QY 389 CTTTTTTCCTCCATTTCTGTGGCAGCTTAATTC-TTTTGTACATCACTTCATCCACCT 447  
Db |||||||  
QY 796 CTTTTTTCCTCCATTTCTGTGGCAGCTTAATTC-TTTTGTACATCACTTCATCCACCT 737  
Db |||||||  
QY 448 TCTGCCATATCAACAGATCCCTTTCTCTATACATCGGCAGCTCATTTATATAGTTGATGT 507  
Db |||||||  
QY 736 TTTGCCATATCAACAGATCCCTTTCTCTATACATCGGCAGCTCATTTATATAGTTGATGT 677  
Db |||||||  
QY 508 TGAATTCAGAAAACAAATCTCATTTCTGTCTGTGNAAGAGTTCCCTGTAATCTCCCTT 567  
Db |||||||  
QY 676 TGAATTCAGAAAACAAATCTCATTTCTGTCTGTGNAAGAGTTCCCTGTAATCTCCCTT 617  
Db |||||||  
QY 568 GGGCTTGACTGTGTTAGTCCAGATTGTTG 598  
Db |||||||  
QY 616 GGGCTTGACTGTGTTAGTCCAGATTGTTG 586  
Db |||||||

RESULT 15  
ABV34731/c  
ID ABV34731 standard; cDNA; 287 BP.  
XX  
AC ABV34731;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 34722.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
WP1; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1; Page 7273; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 287 BP; 97 A; 55 C; 81 G; 54 T; 0 other;  
Query Match 33.9%; Score 202.6; DB 23; Length 287;  
Best Local Similarity 97.0%; Pred. No. 2.3e-52;  
Matches 227; Conservative 0; Mismatches 5; Indels 2; Gaps 2;  
QY 346 AGTGCAAGGGCACTGGCTTGTCTGGTCCGGGTCACTGCCATCTTTTTCCTTCATTT 405  
Db |||||||  
QY 287 AGTGCAAGGGCACTGGCTTGTCTGGTCCGGGTCACTGCCATCTTTTTCCTTCATTT 228  
Db |||||||  
QY 406 CTGTTGGCAGCTTAA-TTCTTTTGTCACTTCAT-CCACCTTCGCATATCAACAC 463  
Db |||||||  
QY 227 CTGTTGGCAGCTTAA-TTCTTTTGTCACTTCAT-CCACCTTCGCATATCAACAC 168  
Db |||||||  
QY 464 AGTCCCTTTCTCATACATCGGCAGCTCATTTATATAGTTGATGTTGAATTCAGAAAACAA 523  
Db |||||||  
QY 167 AGTCCCTTTCTCATACATCGGCAGCTCATTTATATAGTTGATGTTGAATTCAGAAAACAA 108  
Db |||||||  
QY 524 AATCTCATTTTGTCTGTGNAAGAGTTCCCTGTAATCTCCCTTGGGCTTGATC 577  
Db |||||||  
QY 107 AATCTCATTTTGTCTGTGNAAGAGTTCCCTGTAATCTCCCTTGGGCTTGATC 54  
Db |||||||

Search completed: July 12, 2003, 13:44:15  
Job time : 164 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 13:38:58 ; Search time 44 Seconds  
(without alignments)  
4168.016 Million cell updates/sec

Title: US-09-914-191-1  
Perfect score: 598  
Sequence: 1 ttggaatagttctgtcttta.....gggttagtcagattgttg 598

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/1/ina/5A COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PTCUS COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48.4	8.1	7218	1	US-08-232-463-14
2	39.2	6.6	43795	3	US-08-742-185-101
3	35.8	6.0	4084	2	US-08-568-459A-1
4	35.8	6.0	4084	2	US-08-487-826B-1
5	35.8	6.0	4084	4	US-09-210-288-1
6	35.8	6.0	4084	6	5198347-5
7	35.6	6.0	451	2	US-08-943-087-3
8	35.4	5.9	289	4	US-09-007-005-17
9	35.4	5.9	289	4	US-09-244-796-17
10	33.4	5.6	2295	1	US-08-375-300-3
11	33.4	5.6	2295	3	US-09-177-431-3
12	33.4	5.6	2295	5	PCT-US95-16930-3
13	33.4	5.6	4080	1	US-08-375-300-1
14	33.4	5.6	4080	1	US-09-177-431-1
15	33.4	5.6	4080	5	PCT-US95-16930-1
16	33.4	5.6	4705	1	US-07-998-003A-96
17	33.4	5.6	4705	1	US-08-453-274B-96
18	33.4	5.6	4705	1	US-08-453-695A-96
19	33.4	5.6	4705	1	US-08-268-161A-96
20	33.4	5.6	4705	2	US-08-453-702A-96
21	33.4	5.6	4705	4	US-09-099-639-96
22	33.4	5.6	4705	5	PCT-US93-12588-96
23	33.4	5.6	4705	5	PCT-US95-08071-96
24	33.2	5.6	2214	6	5258502-1
25	33	5.5	277	4	US-09-007-005-3
26	33	5.5	277	4	US-09-244-796-3
27	33	5.5	152331	3	US-09-128-155-16

28	33	5.5	176373	3	US-09-128-155-17	Sequence 17, Appl
C 29	32.8	5.5	1055	3	US-09-215-131-3	Sequence 3, Appli
C 30	32.8	5.5	1055	3	US-09-222-734-3	Sequence 3, Appli
C 31	32.8	5.5	2268	2	US-08-890-853-1	Sequence 1, Appli
C 32	32.8	5.5	2268	2	US-09-099-125A-1	Sequence 1, Appli
C 33	32.8	5.5	2268	2	US-09-099-124A-1	Sequence 1, Appli
C 34	32.8	5.5	2268	2	US-09-197-008-1	Sequence 1, Appli
C 35	32.8	5.5	2268	4	US-09-032-476-1	Sequence 1, Appli
C 36	32.8	5.5	2268	4	US-08-890-854-1	Sequence 1, Appli
C 37	32.8	5.5	2268	4	US-09-023-324-1	Sequence 1, Appli
C 38	32.8	5.5	2271	4	US-08-910-820-8	Sequence 8, Appli
C 39	32.8	5.5	2931	4	US-09-168-629-14	Sequence 14, Appl
C 40	32.8	5.5	3966	3	US-09-215-131-1	Sequence 1, Appli
C 41	32.8	5.5	3966	3	US-09-222-734-1	Sequence 32, Appl
C 42	32.6	5.5	248	4	US-09-007-005-32	Sequence 32, Appl
C 43	32.6	5.5	248	4	US-09-244-796-32	Sequence 82, Appl
C 44	32.6	5.5	4145	4	US-09-302-620B-82	Sequence 97, Appl
45	32.4	5.4	787	3	US-08-961-083-97	

RESULT 1

US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZgpt-F1s  
; US-08-232-463-14

Query Match 8.1%; Score 48.4; DB 1; Length 7218;

```
Best Local Similarity 4.1%; Pred. No. 1.5e-05;
Matches 16; Conservative 216; Mismatches 162; Indels 0; Gaps 0;

QY 106 AGAGTAGATACACACCCCATCCATCCCTTCCCTGCTCCCTCCCACTGAGT 165
||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 1057 AGCTGCGATYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1116

QY 166 TGTGTCATTCGACCACTGTCCTGGGTGGTAGGATGCTACAGCCACCTAAGCAAGGAG 225
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1117 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1176

QY 226 CCTGGAGGTGGGAGGCTTGATGTTAAGCACACCAAGTGAAGCCCAAAAGGTC 285
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1177 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1236

QY 286 AGCTGCTTCATCTAGAAATCTGATGTTCTTCAGAAAGATCCCGATGATCGC 345
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1237 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1296

QY 346 AGTGAAGGCACTGGCTTGTCTGCTGGTCCGGTCACTGCATCTTTTCTTCCATTT 405
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1297 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1356

QY 406 CTGTGGCAGCTTAATTTCTTTGTCTCATCTCACTTCACTCTGCTCATCAACACAG 465
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1357 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1416

QY 466 TCCTTTCTCATATACCGGAGCTCATATTATA 499
: : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1417 YVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVYVY 1450

RESULT 2
US-08-742-185-101/c
; Sequence 101, Application US/08742185
; Patent No. 6020476
; GENERAL INFORMATION:
; APPLICANT: Page, David C.
; APPLICANT: Reijo, Renee
; APPLICANT: Saxena, Richa
; APPLICANT: Hawkins, Trevor
; APPLICANT: Reeve, Mary Pat
; TITLE OF INVENTION: DAZ: A GENE FAMILY ASSOCIATED WITH AZOOSPERMIA
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,185
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/690,734
; FILING DATE: 31-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/310,429
; FILING DATE: 22-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI94-07A2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540

; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43795 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-742-185-101

Query Match 6.6%; Score 39.2; DB 3; Length 43795;
Best Local Similarity 53.2%; Pred. No. 0.048;
Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 362 CTTTGTCTGTCGCGGTCACTGCCATCTTTTCTCTCCATTTCTGTTGGCAGCTTAAT 421
||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 38974 CTCTTCTCTTCCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 38915

QY 422 TTTCTTTTGTCACTATTCATCCACCTTCGTCATATCAACACAGTCCCTTTTCTATACAT 481
||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 38914 TTTCTTTTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 38855

QY 482 CGCGAGCTCATTTATATAGTTGATGTTGAATTCAGA 517
||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 38854 CTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 38819

RESULT 3
US-08-568-459A-1/c
; Sequence 1, Application US/08568459A
; Patent No. 5849306
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobb Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/568,459A
; FILING DATE: 07-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium vivax
```

US-08-568-459A-1

Query Match 6.0%; Score 35.8; DB 2; Length 4084;  
Best Local Similarity 53.1%; Pred. No. 0.15;  
Matches 76; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 391 TTTTTCCTTCATTTCTGTGGCAGCTTAATTTCTTTGTGCATCATCTTCACCTTCT 450  
|||||  
Db 1616 TTTTTCATCATATTTTCTTTCAGTTTGTGCAATCTCTGCGCAATCTGCACAGTAAT 1557  
|||||

QY 451 GCCATATCAACACAGTCCCTTCTCTATACATCGGCAGCTCATTTATATTAGTTGATGTTGA 510  
|||||  
Db 1556 CCCTTCCCATCTCGAATCCATCATATATCTCGGGTCTATATTTACCGCAACATTTA 1497  
|||||

QY 511 ATTCAGAAAACAAAATCTCATTC 533  
|||||  
Db 1496 ATTTACAAATCCATATAAAATTC 1474  
|||||

## RESULT 4

US-08-487-826B-1/c  
; Sequence 1, Application US/08487826B  
; Patent No. 5993827  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellens, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,826B  
; FILING DATE: 10-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: NIH121.001CP1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4084 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium vivax  
; US-08-487-826B-1

Query Match 6.0%; Score 35.8; DB 2; Length 4084;  
Best Local Similarity 53.1%; Pred. No. 0.15;  
Matches 76; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 391 TTTTTCCTTCATTTCTGTGGCAGCTTAATTTCTTTGTGCATCATCTTCACCTTCT 450  
|||||

Db 1616 TTTTTCATCATATTTTCTTTCAGTTTGTGCAATCTCTGCGCAATCTGCACAGTAAT 1557  
|||||

QY 451 GCCATATCAACACAGTCCCTTCTCTATACATCGGCAGCTCATTTATATTAGTTGATGTTGA 510  
|||||

Db 1556 CCCTTCCCATCTCGAATCCATCATATATCTCGGGTCTATATTTACCGCAACATTTA 1497  
|||||

QY 511 ATTCAGAAAACAAAATCTCATTC 533  
|||||

Db 1496 ATTTACAAATCCATATAAAATTC 1474  
|||||

## RESULT 5

US-09-210-288-1/c  
; Sequence 1, Application US/09210288  
; Patent No. 6392026  
; GENERAL INFORMATION:  
; APPLICANT: Sim, Kim L.  
; APPLICANT: Chitnis, Chetan  
; APPLICANT: Miller, Louis H.  
; APPLICANT: Peterson, David S.  
; APPLICANT: Su, Xin-zhaun  
; APPLICANT: Wellens, Thomas E.  
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX  
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe Martens Olson & Bear  
; STREET: 620 Newport Center Drive 16th Floor  
; CITY: Newport Beach  
; STATE: California  
; COUNTRY: US  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/210,288  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fuller, Michael  
; REGISTRATION NUMBER: 36,516  
; REFERENCE/DOCKET NUMBER: NIH121.1FWDV1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4084 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Plasmodium vivax  
; US-09-210-288-1

Query Match 6.0%; Score 35.8; DB 4; Length 4084;  
Best Local Similarity 53.1%; Pred. No. 0.15;  
Matches 76; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 391 TTTTTCCTTCATTTCTGTGGCAGCTTAATTTCTTTGTGCATCATCTTCACCTTCT 450  
|||||

Db 1616 TTTTTCATCATATTTTCTTTCAGTTTGTGCAATCTCTGCGCAATCTGCACAGTAAT 1557  
|||||

QY 451 GCCATATCAACACAGTCCCTTCTCTATACATCGGCAGCTCATTTATATTAGTTGATGTTGA 510  
|||||

Db 1556 CCCTTCCCATCTCGAATCCATCATATATCTCGGGTCTATATTTACCGCAACATTTA 1497  
|||||

[illegible]

Qy 478 ACATCGGACGCTCATTTATTAGTTGATGTTGAATTCAGAAAACAAATCTCATCTTGT 537  
Db 114 NYNYSYNY 55  
Qy 538 CTGCTGNAAGAGTCCCTGTAATCTCCCTGGGCTGTGACTGGGTAGTCC 589  
Db 54 AYTYYGYTAYATYYGYTAYATYAYATYAYGYTAYATYYGYTYCYCY 3

## RESULT 9

US-09-244-796-17/c  
; Sequence 17, Application US/09244796  
; Patent No. 6281344  
; GENERAL INFORMATION:  
; APPLICANT: Sztostak, Jack W.  
; APPLICANT: Roberts, Richard W.  
; APPLICANT: Liu, Rihe  
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN  
; FILE REFERENCE: 00786/350007  
; CURRENT APPLICATION NUMBER: US/09/244.796  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/035,963  
; EARLIER FILING DATE: 1997-01-27  
; EARLIER APPLICATION NUMBER: 60/064,491  
; EARLIER FILING DATE: 1997-11-06  
; EARLIER APPLICATION NUMBER: 09/007,005  
; EARLIER FILING DATE: 1998-01-14  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17  
; LENGTH: 289  
; TYPE: RNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Translation template  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(289)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-244-796-17

Query Match 5.9%; Score 35.4; DB 4; Length 289;  
Best Local Similarity 7.8%; Pred. NO. 0.043;  
Matches 18; Conservative 94; Mismatches 120; Indels 0; Gaps 0;

Qy 358 CTGCTTTGCTCGGTCGGGTCACTGCCATCTTTTCTCTCATTTCTGTTGGCAGCT 417  
Db 234 CYGYAYGYCYTYGYSYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNYNY 175  
Qy 418 TAATTTCTTTTGTCTACCTTCATCCACCTTCTGCCATATCAACACAGTCCCTTCTCAT 477  
Db 174 NYNYSYNY 115  
Qy 478 ACATCGGACGCTCATTTATTAGTTGATGTTGAATTCAGAAAACAAATCTCATCTTGT 537  
Db 114 NYNYSYNY 55  
Qy 538 CTGCTGNAAGAGTCCCTGTAATCTCCCTGGGCTGTGACTGGGTAGTCC 589  
Db 54 AYTYYGYTAYATYYGYTAYATYAYATYAYGYTAYATYYGYTYCYCY 3

## RESULT 10

US-08-375-300-3/c  
; Sequence 3, Application US/08375300  
; Patent No. 5679566  
; GENERAL INFORMATION:  
; APPLICANT: Feng, He  
; APPLICANT: Jacobson, Allan S.  
; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN  
; TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street Suite 3100  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/375,300  
; FILING DATE: 20-JAN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Faese, J. P.  
; REGISTRATION NUMBER: 32,983  
; REFERENCE/DOCKET NUMBER: 04020/046001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)542-5070  
; TELEFAX: (617)542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2295 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; US-08-375-300-3  
Query Match 5.6%; Score 33.4; DB 1; Length 2295;  
Best Local Similarity 52.5%; Pred. No. 0.68;  
Matches 73; Conservative 0; Mismatches 66; Indels 0; Gaps 0;  
Qy 350 CAAGGCACTGGCTTGTCTCGTGGTCACTGCCATCTTTTCTCTCATCTTGT 409  
Db 1855 CAAGATCACCACCACTCCAAATCAGAGTCGCTGCTCTTCTTCTTCTCATCAT 1796  
Qy 410 TGGCAGCTTAATTTCTTTTGTCTCATCTTCACCTTCCCATATCAACACAGTCCC 469  
Db 1795 CTTTCATCATCATCTTCTTCTCATCATCTTCATCTTCATCTTCTTCTTCTCATCAT 1736  
Qy 470 TTTCTTATACATCGGCAGC 488  
Db 1735 CGTCATCGTCATCGTCATC 1717  
RESULT 11  
US-09-177-431-3/c  
; Sequence 3, Application US/09177431  
; Patent No. 6071700  
; GENERAL INFORMATION:  
; APPLICANT: He, Feng  
; APPLICANT: Jacobson, Allan S.  
; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE  
; TITLE OF INVENTION: ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/177,431

;; FILING DATE:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/955,472  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fasse, J. Peter  
;; REGISTRATION NUMBER: 32,983  
;; REFERENCE/DOCKET NUMBER: 07917/050001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617/542-5070  
;; TELEFAX: 617/542-9806  
;; TELEX: 200154  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2295 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: cdna  
US-09-177-431-3

Query Match 5.6%; Score 33.4; DB 3; Length 2295;  
Best Local Similarity 52.5%; Pred. No. 0.68;  
Matches 73; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 350 CAAGGCACTGGCTTGTCTGCTGGTCCGGGTCACTGCCATCTTTTTCCTTCCATTTCTGT 409  
DB 1855 CAAGATCACCACCATACTCCAAATCAGAGTGGTCTTCTTCTTCCATCATCATCAT 1796  
QY 410 TGGCAGCTTAATTTCTTTTGTCTCATCTTCCACCTTCTGCGCATATCAACACAGTCCC 469  
DB 1795 CCTCATCATCATCATCTTCATCATCTCCATCGGCTTCTTCCATCATCATCATCAT 1736  
QY 470 TTTCCTATACATCGGCAGC 488  
DB 1735 CGTCATCGTCATCGTCATC 1717

RESULT 12  
PCT-US95-16930-3/c  
;; Sequence 3, Application PC/TUS9516930  
;; GENERAL INFORMATION:  
;; APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL  
;; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE  
;; TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF  
;; TITLE OF INVENTION: NONSENSE-MEDIATED mRNA DECAY  
;; FUNCTION  
;; NUMBER OF SEQUENCES: 6  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Richardson  
;; STREET: 225 Franklin Street Suite 3100  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02110-2804  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US95/16930  
;; FILING DATE: 27-DEC-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/375,300  
;; FILING DATE: 20-JAN-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fasse, J. Peter  
;; REGISTRATION NUMBER: 32,983  
;; REFERENCE/DOCKET NUMBER: 04020/046W01  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617)542-5070  
;; TELEFAX: (617)542-8906

;; TELEX: 200154  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2295 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA  
PCT-US95-16930-3

Query Match 5.6%; Score 33.4; DB 5; Length 2295;  
Best Local Similarity 52.5%; Pred. No. 0.68;  
Matches 73; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 350 CAAGGCACTGGCTTGTCTGCTGGTCCGGGTCACTGCCATCTTTTTCCTTCCATTTCTGT 409  
DB 1855 CAAGATCACCACCATACTCCAAATCAGAGTGGTCTTCTTCTTCCATCATCATCAT 1796  
QY 410 TGGCAGCTTAATTTCTTTTGTCTCATCTTCCACCTTCTGCGCATATCAACACAGTCCC 469  
DB 1795 CCTCATCATCATCATCTTCATCATCTCCATCGGCTTCTTCCATCATCATCATCAT 1736  
QY 470 TTTCCTATACATCGGCAGC 488  
DB 1735 CGTCATCGTCATCGTCATC 1717

RESULT 13  
US-08-375-300-1/c  
;; Sequence 1, Application US/08375300  
;; Patent No. 5679566  
;; GENERAL INFORMATION:  
;; APPLICANT: Feng, He  
;; APPLICANT: Jacobson, Allan S.  
;; TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN  
;; TITLE OF INVENTION: THE ABSENCE OF NONSENSE-MEDIATED mRNA DECAY FUNCTION  
;; NUMBER OF SEQUENCES: 6  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Richardson  
;; STREET: 225 Franklin Street Suite 3100  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02110-2804  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/375,300  
;; FILING DATE: 20-JAN-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Fasse, J. P.  
;; REGISTRATION NUMBER: 32,983  
;; REFERENCE/DOCKET NUMBER: 04020/046001  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617)542-5070  
;; TELEFAX: (617)542-8906  
;; TELEX: 200154  
;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 4080 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA  
US-08-375-300-1

Query Match 5.6%; Score 33.4; DB 1; Length 4080;  
Best Local Similarity 52.5%; Pred. No. 0.96;  
Matches 73; Conservative 0; Mismatches 66; Indels 0; Gaps 0;



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2003, 13:41:28 ; Search time 114 Seconds  
(without alignments)  
8282.800 Million cell updates/sec

Title: US-09-914-191-1

Perfect score: 598

Sequence: 1 ttggaatagttcttctgctta.....ggtgttagtcagattgttg 598

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1105431 seqs, 789497651 residues

Total number of hits satisfying chosen parameters: 2210862

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	597	99.8	1231	9	US-10-098-841-171
C 2	594.4	99.4	1512	9	US-10-198-846-10017
C 3	541.6	90.6	557	9	US-10-106-698-968
C 4	38	6.4	750	9	US-10-184-644-104
C 5	38	6.4	750	9	US-10-184-634-104
C 6	36.6	6.1	104	10	US-09-864-761-31778
C 7	36.6	6.1	597	10	US-09-864-761-15256
C 8	36.2	6.1	734	9	US-10-106-698-3112
C 9	36.2	6.1	2015	9	US-10-084-817-297
C 10	35.8	6.0	4084	9	US-10-153-273-1
C 11	35.8	6.0	185695	9	US-10-020-141-11
C 12	35.8	6.0	185695	9	US-10-017-721-1
C 13	35.4	5.9	432	9	US-09-918-995-8420
C 14	35.2	5.9	594	9	US-10-123-155-10
C 15	35.2	5.9	1919	9	US-09-938-842A-2718
C 16	35	5.9	6660	10	US-09-764-877-2529
C 17	33.6	5.6	401	10	US-09-960-352-7651
C 18	33.4	5.6	2239	10	US-09-925-301-227
C 19	33.4	5.6	2775	9	US-10-125-540-194

C 20	33.4	5.6	2775	10	US-09-764-870-194
C 21	33.2	5.6	264	9	US-10-083-357-80
C 22	33	5.5	4381	9	US-10-177-293-227
C 23	33	5.5	152331	9	US-10-095-407-16
C 24	33	5.5	176373	9	US-10-095-407-17
C 25	32.8	5.5	429	10	US-09-864-761-26138
C 26	32.8	5.5	566	10	US-09-864-761-9764
C 27	32.8	5.5	866	9	US-10-198-846-7248
C 28	32.8	5.5	1558	9	US-10-106-698-455
C 29	32.8	5.5	2268	9	US-10-156-610-1
C 30	32.8	5.5	2268	9	US-09-847-946A-1
C 31	32.8	5.5	2268	9	US-10-243-408-1
C 32	32.8	5.5	2271	9	US-09-844-988-8
C 33	32.8	5.5	2271	9	US-10-338-462-8
C 34	32.8	5.5	2271	10	US-09-844-908-8
C 35	32.8	5.5	2616	10	US-09-771-161A-50
C 36	32.8	5.5	2931	10	US-09-796-872-14
C 37	32.8	5.5	19820	10	US-09-764-877-2713
C 38	32.6	5.5	4145	9	US-10-138-838-82
C 39	32.6	5.5	4145	9	US-10-139-031-82
C 40	32.6	5.5	4145	9	US-10-138-905-82
C 41	32.6	5.5	4145	9	US-10-138-916-82
C 42	32.6	5.5	4145	9	US-09-976-800-82
C 43	32.6	5.5	4145	10	US-09-911-781-3
C 44	32.6	5.5	18996	9	US-10-270-333-61
C 45	32.6	5.5	1691139	9	US-10-067-514-1

## ALIGNMENTS

### RESULT 1

US-10-098-841-171/c  
; Sequence 171, Application US/10098841  
; Publication No US20020197679A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Xu, Chongjun  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Zhang, Jie  
; APPLICANT: Qian, Xiaohong B.  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 784CIP2  
; CURRENT APPLICATION NUMBER: US/10/098,841  
; CURRENT FILING DATE: 2002-03-13  
; PRIOR APPLICATION NUMBER: 09/598,042  
; PRIOR FILING DATE: 2000-06-20  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 331  
; SOFTWARE: pt\_fl\_genes Version 1.0  
; SEQ ID NO 171  
; LENGTH: 1231  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (42)..(938)  
US-10-098-841-171

Sequence 194, Appl  
Sequence 80, Appl  
Sequence 227, Appl  
Sequence 16, Appl  
Sequence 17, Appl  
Sequence 26138, A  
Sequence 9764, Ap  
Sequence 7248, Ap  
Sequence 455, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 8, Appl  
Sequence 8, Appl  
Sequence 50, Appl  
Sequence 14, Appl  
Sequence 2713, Ap  
Sequence 82, Appl  
Sequence 82, Appl  
Sequence 82, Appl  
Sequence 82, Appl  
Sequence 82, Appl  
Sequence 3, Appl  
Sequence 61, Appl  
Sequence 1, Appl

Query Match 99.8%; Score 597; DB 9; Length 1231;  
Best Local Similarity 99.8%; Pred. No. 8.3e-178; Indels 0; Gaps 0;  
Matches 597; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTGAATAGTTCTTGGCTTTATATAAATAGTACTCGATTAAAAAAGACACTTCTGCCA 60  
Db 1221 TTGAATAGTTCTTGGCTTTATATAAATAGTACTCGATTAAAAAAGACACTTCTGCCA 1162

Qy 61 AAGAAACCATGTTCCAAACACCGCAAAAGAGTGTCTGCTTAAACAGAGTAAGATACAC 120  
Db 1161 AAGAAACCATGTTCCAAACACCGCAAAAGAGTGTCTGCTTAAACAGAGTAAGATACAC 1102

Qy 121 ACCCCCATCCATCCCTTCCCTGTTCCCTCCCACTTCCCTGAGTGTCTGCTTAAACAGAGTAAGATACAC 180  
Db 1101 ACCCCCATCCATCCCTTCCCTGTTCCCTCCCACTTCCCTGAGTGTCTGCTTAAACAGAGTAAGATACAC 1042

Qy 181 AGTGTCTGGGTAGGATGCTACAGCCACCTTAAGCAAGGAGCCCTGGGAGTGGGA 240  
Db 1041 AGTGTCTGGGTAGGATGCTACAGCCACCTTAAGCAAGGAGCCCTGGGAGTGGGA 982

Qy 241 GGGCTTGATGTTAAGCACAACAGAACTGAAGCGCAAAAGGTCAGTGTCTTCATCTA 300  
Db 981 GGGCTTGATGTTAAGCACAACAGAACTGAAGCGCAAAAGGTCAGTGTCTTCATCTA 922

Qy 301 GAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCCGAGTCAAGAGGCACTG 360  
Db 921 GAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCCGAGTCAAGAGGCACTG 862

Qy 361 GCTTTGCTGCTGGGTGCTACTGCCATCTTTTTCCTTCCCTGAGTGTCTGAGTGTAA 420  
Db 861 GCTTTGCTGCTGGGTGCTACTGCCATCTTTTTCCTTCCCTGAGTGTCTGAGTGTAA 802

Qy 421 TTTCTTTGTCATCACTTATCCACCTTCTGCCATATCAACAGTCCCTTCTCTATACA 480  
Db 801 TTTCTTTGTCATCACTTATCCACCTTCTGCCATATCAACAGTCCCTTCTCTATACA 742

Qy 481 TCGCAGCTCAATTATAGTGTGATGTTGAATTCAGAAACAAATCTCATTTCTGCTG 540  
Db 741 TCGCAGCTCAATTATAGTGTGATGTTGAATTCAGAAACAAATCTCATTTCTGCTG 682

Qy 541 CTGNAAGATTCCTGTAATCTCCCTGGGCTTGTGCTGTTAGTCCAGATTGTTG 598  
Db 681 CTGNAAGATTCCTGTAATCTCCCTGGGCTTGTGCTGTTAGTCCAGATTGTTG 624

RESULT 2  
US-10-198-846-10017/c  
; Sequence 968, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10017  
; LENGTH: 1512  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1492, 1493, 1494, 1495, 1496, 1497, 1498, 1499, 1500, 1501,  
; LOCATION: 1502, 1503, 1504, 1505, 1506, 1507, 1508, 1509, 1510, 1511,

; LOCATION: 1512  
; OTHER INFORMATION: n = A, T, C or G  
US-10-198-846-10017

Query Match 99.4%; Score 594.4; DB 9; Length 1512;  
Best Local Similarity 99.7%; Pred. No. 6.2e-177; Indels 0; Gaps 0;  
Matches 595; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TGAATAGTTCTTGGCTTTATATAAATAGTACTCGATTAAAAAAGCACTTCTGCCAA 61  
Db 1486 TGAATAGTTCTTGGCTTTATATAAATAGTACTCGATTAAAAAAGCACTTCTGCCAA 1427

Qy 62 AGAAACCATGTTCCAAACACCGCAAAAGAGTGTCTGCTTAAACAGAGTAAGATACACCA 121  
Db 1426 AGAAACCATGTTCCAAACACCGCAAAAGAGTGTCTGCTTAAACAGAGTAAGATACACCA 1367

Qy 122 CCCCATCCATCCCTTCCCTGTTCCCTCCCACTTCCCTGAGTGTCTTCCGACCA 181  
Db 1366 CCCCATCCATCCCTTCCCTGTTCCCTCCCACTTCCCTGAGTGTCTTCCGACCA 1307

Qy 182 GTGCTCTGGGTAGGATGCTACAGCCACCTTAAGCAAGGAGCCCTGGGAGTGGAG 241  
Db 1306 GTGCTCTGGGTAGGATGCTACAGCCACCTTAAGCAAGGAGCCCTGGGAGTGGAG 1247

Qy 242 GGCTTGATGTTTAAAGCACAACAGAACTGAAGCGCAAAAGGTCAGTGTCTTCATCTAG 301  
Db 1246 GGCTTGATGTTTAAAGCACAACAGAACTGAAGCGCAAAAGGTCAGTGTCTTCATCTAG 1187

Qy 302 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGAGTGCAGGCACTGG 361  
Db 1186 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGAGTGCAGGCACTGG 1127

Qy 362 CTTTGTCTGGTCCGGTCACTGCCATCTTTTTCCTTCCCTGAGTGTCTTCCGACCTTAAT 421  
Db 1126 CTTTGTCTGGTCCGGTCACTGCCATCTTTTTCCTTCCCTGAGTGTCTTCCGACCTTAAT 1067

Qy 422 TTTCTTTGTCATCACTTATCCACCTTCTGCCATATCAACAGTCCCTTCTCTATACAT 481  
Db 1066 TTTCTTTGTCATCACTTATCCACCTTCTGCCATATCAACAGTCCCTTCTCTATACAT 1007

Qy 482 CGCAGCTCAATTATAGTGTGATGTTGAATTCAGAAACAAATCTCATTTCTGCTGC 541  
Db 1006 CGCAGCTCAATTATAGTGTGATGTTGAATTCAGAAACAAATCTCATTTCTGCTGC 947

Qy 542 TGAAGAGTTCCTGTAATCTCCCTGGGCTTGTGCTGTTAGTCCAGATTGTTG 598  
Db 946 TGAAGAGTTCCTGTAATCTCCCTGGGCTTGTGCTGTTAGTCCAGATTGTTG 890

RESULT 3  
US-10-106-698-968/c  
; Sequence 968, Application US/10106698  
; Publication No. US20030109690A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides  
; FILE REFERENCE: PA005P1  
; CURRENT APPLICATION NUMBER: US/10/106,698  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: PCT/US00/26524  
; PRIOR FILING DATE: 2000-09-28  
; PRIOR APPLICATION NUMBER: US 60/157,137  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: US 60/163,280  
; PRIOR FILING DATE: 1999-11-03  
; NUMBER OF SEQ ID NOS: 8564  
; SOFTWARE: PatentIn Ver. 3.0  
; SEQ ID NO 968  
; LENGTH: 557  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-106-698-968

Query Match	90.6%;	Score 541.6;	DB 9;	Length 557;
Best Local Similarity	99.8%;	Pred. NO. 1.6e-160;		
Matches 541;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;

  

QY	2	TGGAATAGTCTTCGCTTATATAAATAGTACTGCGATTAAAAAAGACACTTCGCGCAA	61
DB	542	:     :     :     :     :     :     :     :     :     :	
QY	62	AGGAACCATGTTTCCAAACACCGCAAACTGAAGTGTCTGCTTAAACAGAGTAAAGATACACCA	121
DB	482	AGGAACCATGTTTCCAAACACCGCAAACTGAAGTGTCTGCTTAAACAGAGTAAAGATACACCA	423
QY	122	CCCCATCCATCCCTTCCTTCCTGTTCCCTCCCTCCCACTTGAGTTGTGTCATTGCGACCA	181
DB	422	CCCCATCCATCCCTTCCTTCCTGTTCCCTCCCTCCCACTTGAGTTGTGTCATTGCGACCA	363
QY	182	GTGTCCTGGGTGAGGATGCTACAGCCACCTAAGCAAGGAGCCCTGGGAGGTGGGAG	241
DB	362	GTGTCCTGGGTGAGGATGCTACAGCCACCTAAGCAAGGAGCCCTGGGAGGTGGGAG	303
QY	242	GGCTTGATGTTTAAAGCACACCGAACTGAAGCGCAAAAGGGTCAGCTGTCTTCATCTAG	301
DB	302	GGCTTGATGTTTAAAGCACACCGAACTGAAGCGCAAAAGGGTCAGCTGTCTTCATCTAG	243
QY	302	AATCTCGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGGGCAGCTGG	361
DB	242	AATCTCGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGGGCAGCTGG	183
QY	362	CTTTGTCTGTCGGGTCACTGCCATCTTTTTCCTTCCATTTCTGTTGGCAGCTTAAT	421
DB	182	CTTTGTCTGTCGGGTCACTGCCATCTTTTTCCTTCCATTTCTGTTGGCAGCTTAAT	123
QY	422	TTCTTTTGTCTATCACTTCATCCACTTCTGCGCATATCAACACAGTCCCTTTTCCATATCAT	481
DB	122	TTCTTTTGTCTATCACTTCATCCACTTCTGCGCATATCAACACAGTCCCTTTTCCATATCAT	63
QY	482	CGGCAGCTCATATTATATGTTGATGTTGAATTCAGAAAACAAAATCTCATTTCTGTCTGC	541
DB	62	CGGCAGCTCATATTATATGTTGATGTTGAATTCAGAAAACAAAATCTCATTTCTGTCTGC	3
QY	542	TG 543	
DB	2	TG 1	

  

```

RESULT 4
US-10-184-644-104/c
; Sequence 104, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 104
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-104

```

Query Match	6.4%;	Score 38;	DB 9;	Length 750;	
Best Local Similarity	9.0%;	Pred. No. 0.18;			
Matches	39;	Conservative 132;	Mismatches 263;	Indels 0;	Gaps 0;

  

QY	132	TCCCTTCCTCCCTGTTCCCTCCCAACTTGAGTTGTGTGATTCGACCAAGTGTCTGGG	191
DB	683	H..C..H..TY...K..HNKKY..B..NSMH..H..Y..S..MST..A..NMETS...SH..SBSRA	624
QY	192	TGTTAGGATGTCTACAGCCACCTAAGCAAGGAGCCCTGGAGGTGGAGGGCTTGCATG	251
DB	623	MK...DMKS..SR..MHRMY..BBTRHYGH...B..SNT...BKCCYB..TBA..DRM..K..HR..M	564
QY	252	GTTAAGCACACCAAGCAACTGAAGCGCAAAAGGTGACGTGCTTCTCATCTAGAAATCTCTGGA	311
DB	563	B..RA..RBSDR...RCS..MNA..WNMARYTCST..C..Y...B...HNCSC..MS..Y..KCS...	504
QY	312	TGTTCTTCAGAAAGCATCCCCCATGATATCGCAGTGCAGAGGCACTGGCTTTGTCTGTG	371
DB	503	S..SNMWS..R..SMC..C..H..SM..MA..NDB..SRK...AGHEY..ARNC...SSHIN..RTBCY..	444
QY	372	GTCGGGTCACTGCCATCTTTTTTCTTCCTCATTTCTGTGCGACGCTTAATTTCTTTTGTG	431
DB	443	...YSN..TW..ASC..C...THWST...AYV..YWC..MM..AC..SYB...DBBTTCS...H..CC	384
QY	432	ATCACTTCACCCCTTCGCCATATCAACACAGTCCCTTTCCTATACATCGGACGTCA	491
DB	383	BWSHYDC...BRYH...BTCY..AC..BNR..YAB..NASD..DKWM..AS..NCA..C..CBNR..BM	324
QY	492	TTATTATAGTTGATGTTGAATTCAGAAAAAATAATCTCATTTCTGTCTGTGNAAGAGTT	551
DB	323	SCWSSH...TSCKM...M..THRRC...DB...S...CBT..T..CYRTR..NT..RC..A...HCT	264
QY	552	CCCTGTAATCTCCC	565
DB	263	CN.N..NCY..BCCC	250

  

```

RESULT 5
US-10-184-634-104/c
; Sequence 104, Application US/10184634
; Publication No. US20030068684A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C217
; CURRENT APPLICATION NUMBER: US/10/184,634
; CURRENT FILING DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 104
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-634-104

```

  

Query Match	6.4%;	Score 38;	DB 9;	Length 750;	
Best Local Similarity	9.0%;	Pred. No. 0.18;			
Matches	39;	Conservative 132;	Mismatches 263;	Indels 0;	Gaps 0;

  

QY	132	TCCCTTCCTTCCTGTTCCCTCCCAACTTGAGTTGTGTGATTCGACCAAGTGTCTGGG	191
----	-----	---	-----



```
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15256
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL139805.2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.6
US-09-864-761-15256

Query Match          6.1%; Score 36.6; DB 10; Length 597;
Best Local Similarity 65.1%; Pred. No. 0.43; Mismatches 0; Gaps 0;
Matches 54; Conservative 0; Indels 29;

Qy 379 TCACGGCATCTTTTCCATTTCTGTGGCAGCTTAATTTCTTTTGTGTCATCACTT 438
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 308 TCATTCAGCATATTACCTCTTCTTTCTTTTACTTTACTTTCTTCATCTTCATCACTT 367
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 439 CATCCACTTCTGCCATATCAAC 461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 368 CATTCACCTTTTCAATTATCCAC 390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-10-106-698-3112
; Sequence 3112, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3112
; LENGTH: 734
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (252)..(252)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (295)..(295)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (313)..(315)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (321)..(321)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (335)..(335)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (337)..(337)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (351)..(351)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (366)..(369)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (419)..(419)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (435)..(435)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (443)..(443)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (449)..(449)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (452)..(452)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (469)..(469)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (474)..(476)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (484)..(484)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (491)..(491)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (526)..(527)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (537)..(537)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (552)..(552)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (564)..(565)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (578)..(578)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (581)..(581)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (587)..(587)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (589)..(589)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (619)..(619)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (640)..(640)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (644)..(644)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (668)..(668)
; OTHER INFORMATION: n equals a,t,g, or c
```



```
; PRIOR APPLICATION NUMBER: US 60/313,097
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/327,485
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 185695
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-020-141-11

Query Match          6.0%; Score 35.8; DB 9; Length 185695;
Best Local Similarity 59.2%; Pred. No. 25;
Matches 61; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 355 GCACCTGGCTTGTCTGCTGGGTCACTGCATCTTTTTCCTTCATTTCTGTTGGCA 414
Db 77847 GAACCTGCTTTGTCAGCATCCCATAGTTTGAATATGTTGTTCCATTTTGTGTC 77788

Qy 415 GCTTAATTTCTTTGTGCATCATCTTCATCCACCTTCTGCCATAT 457
Db 77787 TCAAAATATAGTTTGATTTCTTTTTCCTTTTCCCTCTGACCTAT 77745

RESULT 12
US-10-017-721-1/c
; Sequence 1, Application US/10017721
; Publication No. US20030096248A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; APPLICANT: Daley, George
; APPLICANT: Bolk, Stacey
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
; FILE REFERENCE: MMI-003
; CURRENT APPLICATION NUMBER: US/10/017,721
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/317,033
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/330,248
; PRIOR FILING DATE: 2001-10-17
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 185695
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-721-1

Query Match          6.0%; Score 35.8; DB 9; Length 185695;
Best Local Similarity 59.2%; Pred. No. 25;
Matches 61; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 355 GCACCTGGCTTGTCTGCTGGGTCACTGCATCTTTTTCCTTCATTTCTGTTGGCA 414
Db 77847 GAACCTGCTTTGTCAGCATCCCATAGTTTGAATATGTTGTTCCATTTTGTGTC 77788

Qy 415 GCTTAATTTCTTTGTGCATCATCTTCATCCACCTTCTGCCATAT 457
Db 77787 TCAAAATATAGTTTGATTTCTTTTTCCTTTTCCCTCTGACCTAT 77745

RESULT 13
US-09-918-995-8420/c
; Sequence 8420, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
```

```
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8420
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(432)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8420

Query Match          5.9%; Score 35.4; DB 9; Length 432;
Best Local Similarity 43.1%; Pred. No. 0.85;
Matches 84; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

Qy 331 CCCGATGATATCGCAGTGCAGGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 390
Db 414 CCCNAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 355

Qy 391 TTTTTCCTTCATTTCTGTTGGCAGCTTAATTTCTTTTGTCTCATCTTCACTTCT 450
Db 354 TTTTNTTTTTTTTTTTTNNNTCAAGAAATGTATCATGATTTTATTTAGTCTGCAC 295

Qy 451 GCCATATCAACACAGTCCCTTCTCTATACATCGGAGCTCATTTATTATAGTTGAT 510
Db 294 GATTTTAAACACAGATCTCTTCAAGATAAGTTAACTTCTATATAAATGATTTT 235

Qy 511 ATTCAAGAAAAACAAA 525
Db 234 CTTCAATAAAAAACAA 220

RESULT 14
US-10-123-155-10/c
; Sequence 10, Application US/10123155
; Publication No. US20030068794A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330RIC30
; CURRENT APPLICATION NUMBER: US/10/123,155
; CURRENT FILING DATE: 2002-04-15
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 10
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-123-155-10

Query Match          5.9%; Score 35.2; DB 9; Length 594;
Best Local Similarity 7.9%; Pred. No. 1.2;
Matches 36; Conservative 163; Mismatches 258; Indels 0; Gaps 0;
```

Search completed: July 12, 2003, 14:26:38  
Job time : 117 secs

```
QY 116 ACACACCCCATCCATCCCTTCTCTCCCTGTTCCCTCCCACTTGAGTTGTCATTC 175
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
565 A.MCTT..AMM.M.CBT.STT.YA.M.YT.S.S.SYSYSYS.S.S.SDSYSYA.SY 506
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 176 GCACCAAGTCTCGGTGGTAGGATGCTACAGCCACCTAAGCAAGAGGCCCTGGGAGG 235
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
505 SYS.S.SWSSYSYSSDDY.CYCYRYHCHSDSYSYY.CRCYYT.SYRYDCHYSC 446
QY 236 TGGAGGGCTTGATGGTTAAGCACACCAAGAACTGAAGCCAAAGGTCAGCTGCTTC 295
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
445 CCSDYVCYSYSYSYSYSYSYSYTDYCSYRCCYYSYSSYSYSSYSYSSAYST 386
QY 296 ATCTAGATCTCGGATCTCTCCAGAAAGCATCCCGATGATCGCAGTCAAGGG 355
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
385 SSSSSSSSYTSTNYC.T.CC...T.MCAABCSTTTTTTTT.HSCC.SA..A. 326
QY 356 CACTGGCTTGTCTGTCGGGTCACTGCATCTTTTCTTCATTTCTGTTGCCAG 415
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
325 M..YC.A.SYSYS.SSS.S.SYMR.HRA.SHYTRS..S.MYCY.YM.Y.YY.YYSY 266
QY 416 CTTAATTTCTTGTCACTCACTCACTCTGCGCATATCAACACAGTCCCTTCTC 475
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
265 YCSRKT...TMDM.T.T..MHY.KYB.HCHTKCRAAT.MN.HTB...N..HBHB.B 206
QY 476 ATACATCGGAGCTATTATATAGTTGATGTTGAATTCAGAAACAAATCTCATCT 535
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
205 H.H..B.H.HSNS...TTS....M.TTM.B.TEASH.RARC.STMM.T.MMM.H.S.Y. 146
QY 536 GTCTGTGNAAGAGTCCCTGTAATCTCCCTGGGCT 572
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
145 .CCR..H.R..R.G...S.CT.HN.B.CYRY.RNGMY 109
```

## RESULT 15

```
US-09-938-842A-2718/c
: Sequence 2718, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kreps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: FILE OF INVENTION: SAME, AND METHODS OF USE
: CURRENT APPLICATION NUMBER: US/09/938,842A
: PRIOR FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/300,111
: PRIOR FILING DATE: 2001-06-22
: NUMBER OF SEQ ID NOS: 5379
: SEQ ID NO 2718
: LENGTH: 1919
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-09-938-842A-2718
```

```
Query Match 5.9%; Score 35.2; DB 9; Length 1919;
Best Local Similarity 60.4%; Pred. No. 2.4; 38; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 3 GGAATAGTCTTCTGCTTTATAAAATAGTACTGCGATTAAAAAAGACACTCTGCGCAA 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1502 GAATAGTCTATCTCTTAAAGAAATGATTGAGATTAAACATGCGATTCCGGTGAA 1443

QY 63 GGAACCATGTTCCCAACCCCAACAGGTGTTCTG 98
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1442 GAATATATTTTAAAGATCCAAATAAACACAGATCAG 1407
```





```

/clone.lib="Soares_NSF_F8_9W_OT_PA_P_S1"
/lab.host="DH10B"
/notes="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HP8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bernaldo."
BASE COUNT      139 a 147 c 111 g 163 t
ORIGIN

```

```

Query Match      92.8%; Score 555; DB 12; Length 560;
Best Local Similarity 99.8%; Pred. No. 3.7e-142;
Matches 555; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTGGAATAGTTCTTGTCTTATAAATAAGTACTGCGATTAAAAAAGACACTTCTGCCA 60
Db 5 TTGGAATAGTTCTTGTCTTATAAATAAGTACTGCGATTAAAAAAGACACTTCTGCCA 64

QY 61 AAGGAACCATGTTTCCAAACACCGCAACAGGTGTTCTGCTTAAACAGAGTAAGATACACC 120
Db 65 AAGGAACCATGTTTCCAAACACCGCAACAGGTGTTCTGCTTAAACAGAGTAAGATACACC 124

QY 121 ACCCCGATCATCCCTTCCCTTCCCTGTTCCCTCCCAACTTGAGTTGTGTCATTCGCACC 180
Db 125 ACCCCGATCATCCCTTCCCTTCCCTGTTCCCTCCCAACTTGAGTTGTGTCATTCGCACC 184

QY 181 AGTGCTCTGGTGTAGGGATGCTACAGCCACCTAAGGCAAGGAGCCCTGGGAGGTGGGA 240
Db 185 AGTGCTCTGGTGTAGGGATGCTACAGCCACCTAAGGCAAGGAGCCCTGGGAGGTGGGA 244

QY 241 GGGCTTCATGTTTAAAGCACACCAAGAACTGAAGCGCAAAAGGGTCAGCTGCTTCACTTA 300
Db 245 GGGCTTCATGTTTAAAGCACACCAAGAACTGAAGCGCAAAAGGGTCAGCTGCTTCACTTA 304

QY 301 GAATCTCTGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGGGCACTG 360
Db 305 GAATCTCTGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGGGCACTG 364

QY 361 GCTTTGCTCTGGTCCGGGTCACTGCCATCTTTTCTCTCCATTTCTGTTGGCAGCTTAA 420
Db 365 GCTTTGCTCTGGTCCGGGTCACTGCCATCTTTTCTCTCCATTTCTGTTGGCAGCTTAA 424

QY 421 TTTCTTTGTCATCATCTTCACTTCCACCTTCTGCCATATCAACACAGTCCCTTTTCTTATACA 480
Db 425 TTTCTTTGTCATCATCTTCACTTCCACCTTCTGCCATATCAACACAGTCCCTTTTCTTATACA 484

QY 481 TCGGCAGCTCATTTATATAGTTGATGTTGAATTCAGAAAAAATAATCTCATTTCTGTCTG 540
Db 485 TCGGCAGCTCATTTATATAGTTGATGTTGAATTCAGAAAAAATAATCTCATTTCTGTCTG 544

QY 541 CTGNAAGAGTTCCCTG 556
Db 545 CTGCAAGAGTTCCCTG 560

```

```

RESULT 2
AL563883 874 bp mRNA linear EST:16-FEB-2001
LOCUS AL563883 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD003XJ17 3
DEFINITION prime, mRNA sequence.
ACCESSION AL563883
VERSION AL563883.1 GI:12913714

```

```

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

```

EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 874)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1..874
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT      206 a 224 c 178 g 252 t 14 others
ORIGIN

```

```

Query Match      83.5%; Score 499.4; DB 9; Length 874;
Best Local Similarity 96.3%; Pred. NO. 7.9e-127;
Matches 498; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 82 GCNAACAAGGTGTTCTGCTTAAACAGAGTAAGATACACACCCCATCCATCCCTTCCTT 141
Db 1 GCNAACAAGGTGTTCTGCTTAAACAGAGTAAGATACACACCCCATCCATCCCTTCCTT 60

QY 142 CCCTGTTCCCTCCCAACTTGAGTTGTGTCATTGCGACCAAGTGTCTCTGGTGTGAGGAT 201
Db 61 CCCTGTTCCCTCCCAACTTGAGTTGTGTCATTGCGACCAAGTGTCTCTGGTGTGAGGAT 120

QY 202 GCTACAGCCACCTTAAGCGAAGGAGCCCTGGGAGGTGGAGGGCTTGCATGGTTAAGCACA 261
Db 121 GCTACAGCCACCTTAAGCGAAGGAGCCCTGGKTTNNNTTGGCTTGCWTGGTTAAGCACA 180

QY 262 CCAAGACTGAAGCGCAAAAGGGTCAGCTGCTTCACTAGAAATCTCTGGATGTCTCTTCC 321
Db 181 CCAGAACTGAAGCGCAAAAGGGTCAGCTGCTTCACTAGAAATCTCTGGATGTCTCTTCC 240

QY 322 AGAAGCATCCCGATGATATCGCAGTGAAGGGCACTGCTTGTCTGCTGGTCCGGGTCA 381
Db 241 AGAAGCATCCCGATGATATCGCAGTGAAGGGCACTGCTTGTCTGCTGGTCCGGGTCA 300

QY 382 CTGCCATCTTTTCTCTCCATTTCTGTTGGCAGCTTAAATTTCTTTTGTTCATCATTCAT 441
Db 301 CTGCCATCTTTTCTCTCCATTTCTGTTGGCAGCTTAAATTTCTTTTGTTCATCATTCAT 360

QY 442 CCACCTTCTGCCATATCAACACAGTCCCTTCCCTATACATCGCAGCTCATTTATTATAGT 501
Db 361 CCACCTTCTGCCATATCAACACAGTCCCTTCCCTATACATCGCAGCTCATTTATTATAGT 420

QY 502 TGATGTTGAATTCAGAAAAAATAATCTCATTTCTGCTGCTGNAAGAGTTCCCTGTGAATC 561
Db 421 TGATGTTGAATTCAGAAAAAATAATCTCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 562 TCCCTTGGGCTTGTACTGTTGTAGTCCAGATGTTG 598

```

```

481 TCCCTGGGCTTGTAAGTGGTGTAGTCCAGATTGTTG 517
Db

RESULT 3
BI084544
LOCUS 602869101T1 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:5013768 3',
DEFINITION mRNA sequence.
ACCESSION BI084544
VERSION BI084544.1 GI:14502874
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 881)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI820 row: a column: 01
High quality sequence start: 28
High quality sequence stop: 722.
Location/Qualifiers
FEATURES
source
1..881
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5013768"
/clone_lib="NIH_MGC_102"
/lab_host="DH10B (phage-resistant)"
/tissue_type="epidermoid carcinoma, cell line"
/notes="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 212 a 236 c 190 g 243 t
ORIGIN
Query Match 81.2%; Score 485.6; DB 13; Length 881;
Best Local Similarity 98.5%; Pred. No. 5e-123;
Matches 521; Conservative 0; Mismatches 5; Indels 3; Gaps 3;
QY 73 TCAACACCGCAACAAGGTGTTCTGCTTAAACAGAGTAAGATACACACCCCATCCAT 132
Db 23 TTAACACCGCAACAAGGTGTTCTGCTTATACAGAGTAAGATACACACCCCTCATCCAT 82
QY 133 CCCTTCCTCCCTGTTCCCTCCCAACTGAGTTGTGTCATTCGCACAGTGTCTGGT 192
Db 83 CCCTTCCTCCCTGTTCCCTCCCAACTGAGTTGTGTCATTCGCACAGTGTCTGGT 142
QY 193 GGTAGGATGTCACAGCACCTAAGGCAAGGAGCCCTGGAGGTGGGAGGCTTGATGG 252
Db 143 GGTAGGATGTCACAGCACCTAAGGCAAGGAGCCCTGGAGGTGGGAGGCTTGATGG 202
QY 253 TTAAGCACACAGAACTGAAGCGCAAAAGGTCAGTGTCTTCTATCTAGAACTCTCGAT 312
Db 203 TTAAGCACACAGAACTGAAGCGCAAAAGGTCAGTGTCTTCTATCTAGAACTCTCGAT 262
QY 313 GTTCCTTCAGAAAGCATCCCGATGATATCGCAGTGCAGGGGACATGGCTTGTCTCG 372
Db 263 GTTCCTTCAGAAAGCATCCCGATGATATCGCAGTGCAGGGGACATGGCTTGTCTCG 322

```

---

```

373 TCCGGGTCACTGCCATCTT-TTTTCTTCATTCTGTGTGGCAGCTTAATTTCTTTGTC 431
Db |||||
323 TCCGGGTCACTGCCATCTTATTTTCTTCTTCATTCTGTGTGGCAGCTTAATTTCTTTGTC 382
QY 432 ATCACTTCATCCACCTTCTGCCATATCAACACAGTCCCTTTCCTATACATCGCGAGTCA 491
Db |||||
383 ATCACTTCATCCACCTTCTGCCATATCAACACAGTCCCTTTCCTATACATCGCGAGTCA 442
QY 492 TTATT-ATAGTTGATGTT-GAATTGAGAAACAAATCTCATTTCTTGTCTGCTGNAAGAG 549
Db |||||
443 TTATTATAGTTGATGTTGCAATTGAGAAACAAATCTCATTTCTTGTCTGCTGCAAGAG 502
QY 550 TTCCTCTGAATCCCTTGGGCTTGTAAGTGTGTTAGTCCAGATTGTTG 598
Db |||||
503 TTCCTCTGAATCTCCCTTGGGCTTGTAAGTGTGTTAGTCCAGATTGTTG 551

RESULT 4
BE206037 479 bp mRNA linear EST 27-JUN-2000
LOCUS bb55G04.x1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:3010614 3',
DEFINITION mRNA sequence.
ACCESSION BE206037
VERSION BE206037.1 GI:8749435
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 479)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Other ESTs: bb55G04.y1
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
Seq primer: -40UP from Gibco
High quality sequence stop: 439...
Location/Qualifiers
FEATURES
source
1..479
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3010614"
/clone_lib="NIH_MGC_17"
/tissue_type="rhabdomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
Site_2: XhoI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

---

```

BASE COUNT 115 a 131 c 97 g 136 t
ORIGIN
Query Match 79.3%; Score 474; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 7.6e-120;
Matches 474; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTGGAATAGTCTTCTTTATATAAAATAGTACTGCGATTAAAAAAGACACTTCTGCCA 60
Db 6 TTGGAATAGTCTTCTTTATATAAAATAGTACTGCGATTAAAAAAGACACTTCTGCCA 65

```

QY 61 AAGGAACATGTTTCAACACCGCAACAAAGGTGTTCTGCTTAAACAGAGTAAGATACACC 120  
 Db 66 AAGGAACATGTTTCAACACCGCAACAAAGGTGTTCTGCTTAAACAGAGTAAGATACACC 125  
 QY 121 ACCCCCATCCATCCCTTCTTCCCTGTTCCCTGTTCCCTCCCAACTTGAGTTGTGTCATTCGCACC 180  
 Db 126 ACCCCCATCCATCCCTTCTTCCCTGTTCCCTGTTCCCTCCCAACTTGAGTTGTGTCATTCGCACC 185  
 QY 181 AGTGTCTGGTGTAGGATGCTACAGCCACCTTAAGGCAAGGAGCCCTGGAGGTGGGA 240  
 Db 186 AGTGTCTGGTGTAGGATGCTACAGCCACCTTAAGGCAAGGAGCCCTGGAGGTGGGA 245  
 QY 241 GGGCTTCATGTTTAAACACACCAAGAACTGAAGCGCAAAAGGTGCTGCTTTCATCTA 300  
 Db 246 GGGCTTCATGTTTAAACACACCAAGAACTGAAGCGCAAAAGGTGCTGCTTTCATCTA 305  
 QY 301 GAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGAGGCACTG 360  
 Db 306 GAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGAGGCACTG 365  
 QY 361 GCTTTGCTGCTGGTCACTGCCATCTTTTCTTCCATTTCTGTTGGCAGCTTAA 420  
 Db 366 GCTTTGCTGCTGGTCACTGCCATCTTTTCTTCCATTTCTGTTGGCAGCTTAA 425  
 QY 421 TTTCTTTTGTGTCATCTATCCACCTTCTGCCATATCAACACAGTCCCTTTTC 474  
 Db 426 TTTCTTTTGTGTCATCTATCCACCTTCTGCCATATCAACACAGTCCCTTTTC 479

RESULT 5  
 BM987297  
 LOCUS  
 DEFINITION UI-H-CO0-aqn-d-10-0-UI.s1 NCI\_CGAP\_Sub9 Homo sapiens cDNA clone  
 IMAGE: 3104658 3', mRNA sequence.  
 ACCESSION BM987297  
 VERSION  
 KEYWORDS EST.  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-x@mail.nih.gov](mailto:cgaps-x@mail.nih.gov)  
 Tissue Procurement: Dr. Jose Mercuende  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LINL at: <http://image.lnl.gov>  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

FEATURES  
 source  
 1..466  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE: 3104658"  
 /clone\_lib="NCI\_CGAP\_Sub9"  
 /tissue\_type="mixed"  
 /dev\_stage="mixed"  
 /lab\_hosts="DH10B (Life Technologies)"  
 /note="Vector: pT73-Pac (Pharmacia) with a modified  
 polylinker; Site 1: EcoR I; Site 2: Not I; tissues:  
 Cholangic mucosa with Crohns disease, Cholangic mucosa with  
 ulcerative colitis, Fetal thymus, Cervix, Cervical  
 adenocarcinoma, Bladder carcinoma, Brain oligodendrocyte;  
 NCI CGAP Sub9 is a subtracted cDNA library constructed  
 according to Bonaldo, Lennon and Soares, Genome Research,

6.791-806, 1996. First strand cDNA synthesis was primed  
 with an oligo-dT primer containing a Not I site. Double  
 stranded cDNA was ligated to an EcoR I adaptor, digested  
 with Not I, and cloned directionally into pT73-Pac  
 vector. The oligonucleotide used to prime the synthesis of  
 first-strand cDNA contains a library tag sequence that is  
 located between the Not I site and the (dT)18 tail. The  
 sequence tags for this library are CGTC, AACG, GGGCC,  
 GGAAG, TAGC, TAAGC, ATGG, AGACA, ATCAC. For additional  
 information, contact: Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 TAG LIB=UI-H-CO0  
 TAG\_TISSUE=Fetal thymus  
 TAG\_SEQ=AACG"

BASE COUNT 110 a 121 c 95 g 140 t  
 ORIGIN  
 Query Match 74.6%; Score 446; DB 14; Length 466;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-112;  
 Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 TGGAAATAGTCTTGTCTTATATAAATAGTACTGCGATTAAAAAAGACACTTCTGCCAA 61  
 Db 21 TGGAAATAGTCTTGTCTTATATAAATAGTACTGCGATTAAAAAAGACACTTCTGCCAA 80  
 QY 62 AGGAACCATGTTCCAAACACCGCAACAAAGGTGTTCTGCTTAAACAGAGTAAGATACACCA 121  
 Db 81 AGGAACCATGTTCCAAACACCGCAACAAAGGTGTTCTGCTTAAACAGAGTAAGATACACCA 140  
 QY 122 CCCCCATCCATCCCTTCCCTGTTCCCTCCCAACTTGTGTCATTCGCACCA 181  
 Db 141 CCCCCATCCATCCCTTCCCTGTTCCCTCCCAACTTGTGTCATTCGCACCA 200  
 QY 182 GTGTCTCGGTGTAGGATGCTACAGCCACCTTAAGCAAGAGCCCTGGAGGTGGAG 241  
 Db 201 GTGTCTCGGTGTAGGATGCTACAGCCACCTTAAGCAAGAGCCCTGGAGGTGGAG 260  
 QY 242 GGCTTCATGTTTAAACACACCAAGAACTGAAGCGCAAAAGGTGCTGCTTTCATCTAG 301  
 Db 261 GGCTTCATGTTTAAACACACCAAGAACTGAAGCGCAAAAGGTGCTGCTTTCATCTAG 320  
 QY 302 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGAGGCACTGG 361  
 Db 321 AATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTCAAGAGGCACTGG 380  
 QY 362 CTTTGTCTGGTCCGGTCACTGCCATCTTTTCTTCCATTTCTGTTGGCAGCTTAAT 421  
 Db 381 CTTTGTCTGGTCCGGTCACTGCCATCTTTTCTTCCATTTCTGTTGGCAGCTTAAT 440  
 QY 422 TTTCTTTTGTGTCATCTTCCATTCATCCACCT 447  
 Db 441 TTTCTTTTGTGTCATCTTCCATTCATCCACCT 466

RESULT 6  
 AW954708/c  
 LOCUS  
 DEFINITION MAGE resequences, MAGC Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AW954708  
 VERSION AW954708.1  
 KEYWORDS GI:8144391  
 SOURCE EST.  
 ORGANISM human.

REFERENCE  
 AUTHORS Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 TITLE  
 JOURNAL Assessment of gene expression patterns in a model of colon tumor  
 metastasis using a 19,200 element cDNA microarray  
 COMMENT Unpublished (2000)  
 Contact: John Quackenbush  
 The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 3528  
 Fax: 301 838 0208  
 Email: johnq@tigr.org  
 Plate: 67

Seq primer: Reverse.

FEATURES  
 source  
 1. .577  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="MAGE resequencing, MAGC"  
 /note="Vector: pbluescriptSKM"  
 BASE COUNT 142 a 102 c 123 g 99 t 111 others  
 ORIGIN

Query Match 73.7%; Score 440.8; DB 10; Length 577;

Best Local Similarity 96.2%; Pred. No. 1.1e-110; Indels 2; Gaps 1;  
 Matches 457; Conservative 0; Mismatches 16;

Qy 125 CCATCCATCCCTCCCTCCCTGTTCC--CTTCCCAACTGAGTTGTGTTCATTCGCACGAG 182

Db 475 CCANNANNCTTTCTNNCCNTGTCCNCCTCCAACTGAGGTGTGTTCATTCGCACGAG 416

Qy 183 TGTCTGGTGGTAGGATGCTACAGCCACCTAAGGCAAGAGGCCCTGGAGGTGGAGG 242

Db 415 TGTCTGGTGGTAGGATGCTACAGCCACCTAAGGCAAGAGGCCCTGGAGGTGGAGG 356

Qy 243 GTTTCATGGTTAAGCACCAGCAACTGAAGCGCAAAAGGCTAGCTCTTCACTAGA 302

Db 355 GTTTCATGGTTAAGCACCAGCAACTGAAGCGCAAAAGGCTAGCTCTTCACTAGA 296

Qy 303 ATCTCTGGATGTTCTTCCGAAAGCATCCCGATGATATCGCAGTGAAGGGCACTGGC 362

Db 295 ATCTCTGGATGTTCTTCCGAAAGCATCCCGATGATATCGCAGTGAAGGGCACTGGC 236

Qy 363 TTGTCTCTGGTCCGGGTCACTGCCATCTTTTTCCTTCCATTTCTGTGGCAGCTTAATT 422

Db 235 TTGTCTCTGGTCCGGGTCACTGCCATCTTTTTCCTTCCATTTCTGTGGCAGCTTAATT 176

Qy 423 TCTTTTGTCACTTCACTCCACCTCTGCCATATCAACAGTCCCTTTCCCTATATC 482

Db 175 TCTTTTGTCACTTCACTCCACCTCTGCCATATCAACAGTCCCTTTCCCTATATC 116

Qy 483 GGCAGCTCATTTATATAGTTGATGTTGAATTCAGAAAACAAATCTCATTTCTGTCT 542

Db 115 GGCAGCTCATTTATATAGTTGATGTTGAATTCAGAAAACAAATCTCATTTCTGTCT 56

Qy 543 GNAAGAGTTCCTGTATATCTCCCTTGGCTGTGACTGTGTAGTCCAGATTGTT 597

Db 55 GCAAGAGTTCCTGTATATCTCCCTTGGCTGTGACTGTGTAGTCCAGATTGTT 1

RESULT 7  
 AA862635  
 LOCUS  
 DEFINITION oh07e08.s1 NCI\_CGAP\_Kid3 Homo sapiens cDNA clone IMAGE:1457126 3',  
 mRNA sequence.  
 ACCESSION AA862635  
 VERSION AA862635.1 GI:2955114  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
 Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 www-bio.llnl.gov/bbrp/image/image.html  
 Seq primer: -40ml3 fwd. ET from Amersham  
 High quality sequence stop: 421.

FEATURES  
 source

1. .439  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NCI\_CGAP\_Kid3"  
 /lab\_host="DH10B"  
 /note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with  
 a modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer,  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not  
 I and Eco RI sites of the modified pT73 vector. mRNA  
 source: 2 pooled kidneys. Library went through one round  
 of normalization. Library constructed by Bento Soares and  
 M. Fatima Bonaldo."

BASE COUNT 108 a 116 c 95 g 120 t

ORIGIN

Query Match 72.9%; Score 435.8; DB 9; Length 439;

Best Local Similarity 99.5%; Pred. No. 2.5e-109; Indels 0; Gaps 0;  
 Matches 437; Conservative 0; Mismatches 2;

Qy 3 GGAATAGTTCCTGCTTTATAAAATAGTACTGCGATTAAAAAAGCACTTCTGCCAAA 62

Db 1 GGAATAGTTCCTGCTTTATAAAATAGTACTGCGATTAAAAAAGCACTTCTGCCAAA 60

Qy 63 GGAACCATGTTCCAAACCCGCAACCAAGGTGTTCTGCTTAAACAGAGTAAGATACAC 122

Db 61 GGAACCATGTTCCAAACCCGCAACCAAGGTGTTCTGCTTAAACAGAGTAAGATACAC 120

Qy 123 CCCCATCCATCCCTTCTTCCCTGTTCCCTCCCAACTTGAGTTGTGTCATTTCGCAC 182

Db 121 CCCCATCCATCCCTTCTTCCCTGTTCCCTCCCAACTTGAGTTGTGTCATTTCGCAC 180

Qy 183 TGTCTGGTGGTAGGATGCTACAGCCACTAAGGCAAGAGGCCCTGGAGGTGGAGG 242

Db 181 TGTCTGGTGGTAGGATGCTACAGCCACTAAGGCAAGAGGCCCTGGAGGTGGAGG 240

Qy 243 GCTTGCATGTTAAGCACACACAGAACTGAAGGCCAAAAGGGTCAGCTGTTTCATCTAGA 302

Db 241 GCTTGCATGTTAAGCACACACAGAACTGAAGGCCAAAAGGGTCAGCTGTTTCATCTAGA 300

Qy 303 ATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGGCGCACTGGC 362

Db 301 ATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGGCGCACTGGC 360

Qy 363 TTTGTCTGTGTCGGGTCACTGCGCATCTTTTTCCTTCCATTTCTGTGGCAGCTTAATT 422

Db 361 TTTGTCTGTGTCGGGTCACTGCGCATCTTTTTCCTTCCATTTCTGTGGCAGCTTAATT 420

Qy 423 TCTTTTGTCACTCACTTCAT 441

Db 421 TCTTTTGTCACTCACTTCAT 439

RESULT 8

AA308526/c

LOCUS

DEFINITION AA308526

ACCESSION sapiens cDNA 5' end, mRNA sequence.

VERSION AA308526.1 GI:1960854

KEYWORDS EST.

SOURCE human.

AA308526 427 bp mRNA linear EST 18-APR-1997  
 EST179347 HCC cell line (metastasis to liver in mouse) II Homo







/clone lib="OVARC1"  
/tissue type="ovary, tumor tissue"  
/note="Vector: pME185FL3"  
BASE COUNT 186 a 135 c 145 g 149 t 3 others  
ORIGIN

Query Match 60.9%; Score 364.4; DB 9; Length 618;  
Best Local Similarity 98.6%; Pred. No. 1.2e-89;  
Matches 365; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 229 TGGAGGTGGAGGCTTGCATGTTAAGCACACAGAACTGAAGCGCAAAAGGTCAGC 288  
Db 618 TGGNAGGTGGAGGCTTGCATGTTAAGCACACAGAACTGAAGCGCAAAAGGTCAGC 559  
QY 289 TGTCTTCATCTAGAACTCTCGATGTTCCCTTCCAGAAAGCATCCCGATGATCGCAGT 348  
Db 558 TGTCTTCATCTAGAACTCTCGATGTTCCCTTCCAGAAAGCATCCCGATGATCGCAGT 499  
QY 349 GCAAGGACACCTGGCTTGTCTGCTCGGGTCACTGCCATCTTTTCCCTTCCATTTCTG 408  
Db 498 GCAAGGACACCTGGCTTGTCTGCTCGGGTCACTGCCATCTTTTCCCTTCCATTTCTG 439  
QY 409 TTGCAGCTTAATTTCTTTTCTATCACTTCCATCCACCTTCTGCCATATCAACACAGTCC 468  
Db 438 TTGCAGCTTAATTTCTTTTCTATCACTTCCATCCACCTTCTGCCATATCAACACAGTCC 379  
QY 469 CTTTCCATATACATCGGAGCTCATTATATAGTTGATGTTGAATTCAGAAACAAATCT 528  
Db 378 CTTTCCATATACATCGGAGCTCATTATATAGTTGATGTTGAATTCAGAAACAAATCT 319  
QY 529 CATTCTTGTCTGTGNAAGAGTCCCTGTATCTCCCTTGGGCTTGTACTGGTGTAGTC 588  
Db 318 CATTCTTGTCTGTGNAAGAGTCCCTGTATCTCCCTTGGGCTTGTACTGGTGTAGTC 259  
QY 589 CAGATTGTTG 598  
Db 258 CAGATTGTTG 249

RESULT 12  
AL530910/c  
LOCUS  
DEFINITION AL530910 LTI\_NFL001\_NBC4 Homo sapiens cDNA clone CS0DD003YU17 5  
prime, mRNA sequence.  
ACCESSION AL530910  
VERSION AL530910.1 GI:12794403  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 929)  
Li, W.B., Gruber, C., Jesse, J., and Polayes, D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers  
1. .929  
/organism="Homo sapiens"  
/db xref="taxon:9606"  
/clone="CS0DD003YU17"  
/clone\_lib="LTI\_NFL001\_NBC4"  
/sex="male"  
/tissue type="neuroblastoma cells"  
/lab host="DH10B"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-stranded cDNA was digested with Not I and  
cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
vector. Library was normalized. Library was constructed

# FEATURES

source

by Life Technologies. Contact : Feng Liang Life  
Technologies, a division of Invitrogen 9800 Medical Center  
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610  
8371 Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com"  
BASE COUNT 263 a 206 c 239 g 220 t 1 others  
ORIGIN

Query Match 58.5%; Score 349.6; DB 9; Length 929;  
Best Local Similarity 99.2%; Pred. No. 1.5e-85;  
Matches 360; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
QY 237 GGGAGGCTTGCAT-GGTTAAGCACACCAAGCTGAAGCGCAAAAGGTCAGCTGCTTC 295  
Db 929 GGGAGGCTTGCATGGGTAAAGCACACCAAGCTGAAGCGCAAAAGGTCAGCTGCTTC 870  
QY 296 ATCTAGAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGGG 355  
Db 869 ATCTAGAATCTCTGGATGTTCTTCCAGAAAGCATCCCGATGATATCGCAGTGCAGGG 810  
QY 356 CACTGGCTTGTCTGGTCCGGTCACTGCCATCTTTTCCCTTCCATTTCTGTTGGCAG 415  
Db 809 CACTGGCTTGTCTGGTCCGGTCACTGCCATCTTTTCCCTTCCATTTCTGTTGGCAG 750  
QY 416 CTTAATTTCTTTTGTTCATCACTTCCACCTTCTGCCATATCAACACAGTGCCTTTCT 475  
Db 749 CTTAATTTCTTTTGTTCATCACTTCCACCTTCTGCCATATCAACACAGTGCCTTTCT 690  
QY 476 ATACATCGGAGCTCATTATATAGTTGATTTGAATTCAGAAACAAATCTCATTTCT 535  
Db 689 ATACATCGGAGCTCATTATATAGTTGATTTGAATTCAGAAACAAATCTCATTTCT 630  
QY 536 GTCTGCTGNAAGAGTCCCTGTAATCTCCCTTGGGCTTGTACTGGTGTAGTCCAGATTG 595  
Db 629 GTCTGCTGNAAGAGTCCCTGTAATCTCCCTTGGGCTTGTACTGGTGTAGTCCAGATTG 570  
QY 596 TTG 598  
Db 569 TTG 567

RESULT 13  
W93172/c  
LOCUS  
DEFINITION W93172.1 GI:1422325

ACCESSION W93172  
VERSION W93172.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 357)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman  
M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J.,  
Rifkin, D., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston  
R., Williamson, A., Wohldmann, P., and Willson, R.  
The WashU-Merck EST Project  
Unpublished (1995)

TITLE  
JOURNAL  
COMMENT  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@wustl.edu  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 466 Std Error: 0.00  
Seq primer: mob.REGA-ET  
High quality sequence stop: 337.  
Location/Qualifiers

# FEATURES





Search completed: July 12, 2003, 14:23:25  
Job time : 1072 secs

VERSION B0073764.1 GI:19902810  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL 1 (bases 1 to 1044)  
COMMENT NIH-MGC <http://mgi.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L1CM2047 row: m column: 24  
High quality sequence stop: 715.

FEATURES  
source  
1..1044  
Location/Qualifiers  
/organism="Homo sapiens"  
/db xref="taxon:9606"  
/clone="IMAGE:5806655"  
/clone\_lib="NIH\_MGC\_101"  
/tissue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:  
XhoI; cDNA made by oligo-dt priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH MGC Library."

BASE COUNT 250 a 228 c 288 g 236 t 2 others  
ORIGIN

Query Match 55.0%; Score 328.8; DB 14; Length 1044;  
Best Local Similarity 98.2%; Pred. No. 8.3e-80;  
Matches 374; Conservative 0; Mismatches 3; Indels 4; Gaps 4;  
Qy 220 AAGCAGCCCTGGGAGGTGGGAGGCTT-GCATGGTTAAGCACA-CCAGAACTGAAGCGCA 277  
Db 890 AGGAAGCCCTGGGAGGTGGGAGGCTTGGCATGGTTAAGCACA-CCAGAACTGAAGCGCA 831  
Qy 278 AAAGGGTCAGCTGTCTTCATAGAACTCTGGATGTTCTTCCAGAAAGCATCCCGAT 337  
Db 830 AAAGGGTCAGCTGTCTTCATAGAACTCTGGATG-TCTTCCAGAAAGCATCCCGAT 772  
Qy 338 GATATCGCAGTGCAGGCGCATCTGGCTTTGCTCGGTCGGGTCACTGCCATCTTTTTC 397  
Db 771 GATATCGCAGTGCAGGCGCATCTGG-TTTGCTCGGTCGGGTCACTGCCATCTTTTTC 713  
Qy 398 TTCCATTTCGTCGCGAGCTTAATTTCTTTTGTGTCATCACTTCATCCACCTTCGCCATAT 457  
Db 712 TTCCATTTCGTCGCGAGCTTAATTTCTTTTGTGTCATCACTTCATCCACCTTCGCCATAT 653  
Qy 458 CAACACAGTCCCTTTTCCTATACATCGGCGAGCTCAATTATATAGTTGATTTGAATTCAGA 517  
Db 652 CAACACAGTCCCTTTTCCTATACATCGGCGAGCTCAATTATAGTTGATTTGAATTCAGA 593  
Qy 518 AAACAAAATCTCATTTCTGTCGTGNAAGAGTTCCCTGTAAATCTCCCTTGGGCTTGTAC 577  
Db 592 AAACAAAATCTCATTTCTGTCGTGCAAGAGTTCCCTGTAAATCTCCCTTGGGCTTGTAC 533  
Qy 578 TGGTGTAGTCCAGATTGTTG 598  
Db 532 TGGTGTAGTCCAGATTGTTG 512